Trends in Biotechnology

CellPress
REVIEWS

Review

CRISPR/Cas-Mediated Base Editing: Technical Considerations and Practical Applications

Kutubuddin A. Molla[1,2] and Yinong Yang[1,*]

Genome editing with CRISPR/Cas has rapidly gained popularity. Base editing, a new CRISPR/Cas-based approach, can precisely convert one nucleotide to another in DNA or RNA without inducing a double-strand DNA break (DSB). A combination of catalytically impaired nuclease variants with different deaminases has yielded diverse base-editing platforms that aim to address the key limitations such as specificity, protospacer adjacent motif (PAM) compatibility, editing window length, bystander editing, and sequence context preference. Because new base editors significantly reduce unintended editing in the genome, they hold great promise for treating genetic diseases and for developing superior agricultural crops. We review here the development of various base editors, assess their technical advantages and limitations, and discuss their broad applications in basic research, medicine, and agriculture.

Highlights

Base editing represents a new dimension of CRISPR/Cas-mediated precise editing to generate single-nucleotide changes in DNA or RNA independently of double-strand breaks and homology-directed repair. Since its invention in 2016, many base-editing tools have been developed to install point mutations in a diverse array of animal, plant, and microbial organisms.

Base editing yields a high efficiency of editing with very low rates of indel formation. Rapid advances in base-editing techniques have significantly reduced unintended editing and expanded the scope and utility of genome targeting.

Base Editors work in both dividing and non-dividing cells and can be applied to correct 61% of human pathogenic mutations listed in the ClinVar database.

Base editing has drawn great academic and industrial interest because it is broadly applicable to basic research, synthetic biology, therapeutics, and crop improvement.

CRISPR/Cas for Single-Nucleotide Alteration

Precision genome editing, including altering single bases, is a versatile and powerful tool to accelerate gene therapy and crop improvement. Generally, **genome editing** (see Glossary) involves the generation of a site-specific double-strand DNA break (DSB) followed by two main types of repair: NHEJ (non-homologous end-joining) or HDR (homology-directed repair) [1–3]. However, the introduction of a precise mutation by genome editing strongly depends on HDR occurring at a DSB site in a genome via a donor DNA template harboring the desired change [4,5]. Although **CRISPR/Cas9** can easily create a DSB at a specific locus [2], HDR in higher eukaryotes is very inefficient because of the low innate rate of homologous recombination and difficulties in onsite delivery of donor DNA [5,6]. The occurrence of unintended **indels** might also cause imprecise editing of the target gene [7,8]. Recently, CRISPR/Cas-mediated **base-editing** systems have been developed to circumvent these limitations [9–11]. Base editing is a new genome-editing technique that generates mutations at single-base resolution. All four transition mutations, C → T, G → A, A → G, and T → C, can be installed in the genome with the available CRISPR/Cas base editors (BEs). The **cytosine base editor** (CBE) can install a C–G to T–A mutation (Box 1), while the **adenine base editor** (ABE) can alter an A–T base pair into a G–C pair (Box 2). In RNA, conversion of Adenine (A) to Inosine (I) is also possible with the **RNA base editor** (RBE) (Box 2).

Unlike regular CRISPR/Cas-mediated genome-editing techniques, BEs do not create a DSB, and therefore indel generation is limited. As a result, BEs offer precise genome editing with much cleaner product output, reducing on- and off-target indels [12]. The molecular genetics and chemistry behind the development of BEs and their application in sequence diversification and other areas have recently been reviewed [7,13,14]. In this review, we present a comprehensive but

[1]Department of Plant Pathology and Environmental Microbiology, and Huck Institute of the Life Sciences, Pennsylvania State University, University Park, PA 16802, USA
[2]ICAR-National Rice Research Institute, Cuttack 753006, India

*Correspondence:
yuy3@psu.edu (Yinong. Yang).

Box 1. Cytosine Base Editors (CBEs)

CBEs, the first group of BEs, were developed to convert targeted C to T (G to A in the opposite strand) [9,10]. A CBE skeleton contains four different components: a cytosine deaminase that catalyzes the conversion of C to U, a modified Cas9 (nCas9/dCas9) that binds target DNA, a sgRNA which direct Cas9-cytosine deaminase to bind the target locus, and a UGI that subverts the cellular uracil base excision repair (BER) pathway (Figure I). To avoid unnecessary indel formation by DSB creation, CBEs have been constructed using either dCas9 (catalytically inactive Cas9) or nCas9 (D10A nickase). Komor and coworkers linked rat cytosine deaminase (rAPOBEC1) to the N terminus of dCas9/nCas9 via an XTEN linker and UGI to the C terminus [9]. A cytosine deaminase from sea lamprey (PmCDA1) has been employed to construct CBE, where both PmCDA1 and UGI are linked to the C terminus of dCas9/nCas9 [10]. Cytosine deaminase is known to act on single strand DNA (ssDNA) [101]. A small window of ssDNA in the noncomplementary strand, created by sgRNA-Cas9-mediated R-loop formation, is targeted by cytosine deaminase. This fact determines the activity window of BEs, typically bases 4–8 of protospacer where PAM is counted as 21–23 bases [9]. However, PmCDA1 containing target-AID has an editing window slightly shifted relative to BE3, optimally at bases 2–6 of the protospacer [10]. When a genomic C is converted to U by cytosine deaminase, the resulting U–G mismatch is either replicated into T–A or restored to C–G by UDG-mediated BER. Because the two alternative pathways compete, overexpression of an inhibitor of UDG increases the efficiency of base editing [9,10]. When the target C is in a favorable window, a base-editing efficiency of up to 75% has been achieved with BE3, compared with 96.1% with target-AID, much higher than the efficiency of HDR-mediated precision editing [9,10].

Although two initial studies [9,10] demonstrated the efficacy of CBEs in mammalian and yeast cell lines, CBEs have been shown to be effective in other organisms including bacteria, mouse, rabbit, *Xenopus*, *Bombyx mori*, and Zebrafish [16,64,87–90]. Similarly, CBEs have been successfully employed to create targeted base alterations in several plant species including rice, maize, wheat, tomato, potato, and *Arabidopsis* [44,45,76,95,97,102].

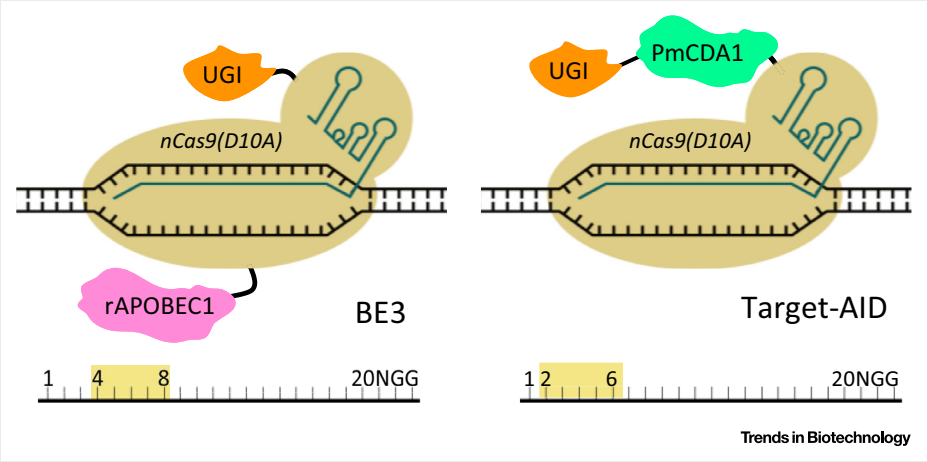


Figure I. Schematic Diagram of Two Original Cytosine Base Editors. BE3: SpCas9 nickase (D10A) is linked to rat cytidine deaminase (rAPOBEC1) (pink) through the N terminus, and to uracil glycosylase inhibitor (UGI) (orange) at the C terminus. Approximate deamination window ranges from bases 4 to 8 of protospacer [when protospacer adjacent motif (PAM) is 21–23]. Target-AID: the C terminus of SpCas9 nickase (D10A) is linked to both cytidine deaminase from *Petromyzon marinus* (PmCDA1) (lime green) and UGI. The estimated activity window of target-AID is from protospacer positions 2 to 6.

concise and updated overview of various base-editing platforms, with critical discussion of their technical advantages and limitations, their distinctness in efficiency and suitability in specific genomic contexts, and their broad applications in basic research and biotechnology.

Precise Base Editing with Reduced Indels and Off-Target Changes

One of the notable advantages of BEs over conventional genome editing via Cas9 [9–11] or Cas12a (also known as Cpf1) [15] is the recovery of a much cleaner edited product with very few to negligible amounts of indels [12] or **off-target editing** [16,17] (Table 1). Unlike CRISPR/Cas-mediated HDR, base editing offers precision editing without double-stranded (ds)DNA

Glossary

Activity window: a range of bases in the protospacer sequence which is favorable for the editing activity of a BE. Activity windows vary across base-editing platforms. Most BEs have an activity window of about 5–6 nt.

Adenine base editor (ABE): composed of catalytically impaired nuclease and laboratory evolved DNA-adenosine deaminase, ABEs convert a targeted A–T base pair to a G–C base pair by deaminating adenosine in the DNA.

Base editing: a CRISPR/Cas-mediated genome-editing method that uses a combination of a catalytically impaired nuclease and a nucleotide deaminase to introduce a point mutation at a target locus in DNA or RNA.

CRISPR/Cas9: clustered regularly interspaced palindromic repeats (CRISPR)/CRISPR-associated protein (Cas9), a two-component genome editing system including a single-guide RNA (sgRNA) and a Cas9 nuclease. A predesigned sgRNA directs Cas9 to bind to and cut a DNA sequence upstream of the PAM.

Cytosine base editor (CBE): composed of a catalytically impaired nuclease and cytosine deaminase, CBEs convert a targeted C–G base pair to a T–A base pair by deaminating cytosine in DNA.

dCas9: catalytically dead Cas9, developed by mutating both the nuclease (RuvC and HNH) domains of Cas9. dCas9 lacks DNA-cleavage activity but retains RNA-directed DNA-binding activity.

Genome editing: a molecular method that makes specific changes in a genome by the deletion, insertion, or replacement of a fragment or specific bases of the genome, which allows the precise removal, addition, or alteration of genetic material.

Indel: insertion (in) or deletion (del) of nucleotides in genomic DNA. Genomic cleavage by Cas9 or other nucleases is followed by indel generation. Indels generally cause frameshift mutations except when the length of the indel is 3 nt or multiple of 3 nt.

nCas9: nickase-Cas9, a mutated form of a Cas9 that creates a nick in

Box 2. Adenine Base Editors (ABEs) and RNA Base Editors (RBEs)

ABEs

ABEs have been developed to convert a targeted A to G (T to C in the opposite strand) [11]. Theoretically, fusing an adenosine deaminase with dCas9/nCas9 would give rise to ABE. However, no enzyme was known to deaminate adenine in DNA [103]. Gaudelli and colleagues carried out extensive directed protein evolution to develop an adenine deaminase that can act on a DNA substrate [11]. *E. coli* tRNA adenosine deaminase (ecTadA) was subjected to evolution, and the inclusion of some mutations resulted in TadA* which can efficiently (53%) deaminate adenine in DNA. Heterodimeric TadA (wtTadA–TadA*) was fused with nCas9 (D10A) to develop the four classes of ABEs – ABE6.3, ABE7.8, ABE7.9, and ABE7.10 (Figure IA). Because the original ecTadA acts on small single-stranded (ss) anticodon loop of tRNA, the ABEs also act on ssDNA in the R-loop formed by binding of sgRNA–nCas9 to the target genomic locus. ABEs favorably deaminate in a window of ~4–6 nt. Among the four classes, ABE7.10 is the most efficient, and prefers to target A at protospacer positions 4–7 (PAM counted as 21–23), whereas the other three perform better when A is at position 8–10. ABEs catalyze the deamination of adenine to inosine, which is treated as guanosine by the polymerase, and, following DNA replication, the A–T base pair is ultimately converted to a G–C base pair. This initial development was followed by several studies demonstrating the application of ABEs in other organisms including mouse, rabbit, rice, wheat, *Arabidopsis*, and *Brassica napus* [23,24,26,27,64,98].

RBEs

RNA editing enables a protein to be altered without any permanent changes in the genome. Before a transcript RNA translates into protein, it can be targeted for specific base editing. Recently, the RNA editing for programmable A to I replacement (REPAIR) system has been developed by combining dCas13 with ADAR2 deaminase (Figure IB) [86]. Because inosine is read as guanosine by the splicing and translation machinery, erroneous G → A mutations could be corrected using the platform. Cas13 can be programmed to bind to specific RNA [104]. Using a 50 nt protospacer with a mismatched cytidine opposite to the target adenine to be edited, authors achieved 28% editing of the disease-relevant mutations tested [86]. Because Cas13 requires no PAM, virtually all RNA is targetable by REPAIRv1. REPAIRv2, an improved version, decreased the off-target editing 900-fold compared with REPAIRv1 [86].

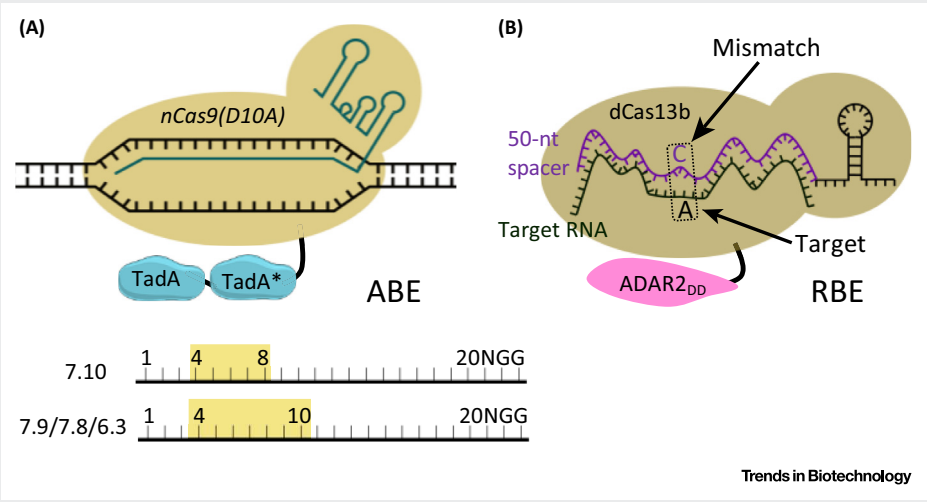


Figure I. Schematic Diagrams of Adenine Base Editor (ABE) and RNA Base Editor (RBE). (A) ABE: fusion of artificially evolved DNA adenine deaminase (TadA–TadA*) (teal) with SpCas9 nickase (D10A) generates ABE (TadA, wild-type *Escherichia coli* tRNA adenosine deaminase; TadA*, mutated TadA). Activity windows are highlighted in the protospacer sequence. Abbreviations: 7.10, ABE7.10; 6.3, ABE6.3; 7.9, ABE7.9; 7.8, ABE7.8. (B) RBE: catalytically dead *Prevotella* sp. Cas13 (dCas13b) is tethered with deaminase domain of human 'adenosine deaminase acting on RNA' (ADAR2DD) to develop RBE. It uses a 50 nt guide with a mismatched cytidine opposite to the target adenosine. Hybridization of the guide (violet) with the target RNA (black) forms the dsRNA required for ADAR2DD action.

cleavage because it is either based on **dCas9** or **nCas9**. Once a genomic DSB is detected, NHEJ competes with HDR for DNA repair, often leading to more indel formation than gene replacement [18]. As a result, correction of mutations using HDR is very inefficient [19].

In contrast to cytosine base editing, that has 37% efficiency and generates 1.1% indels, HDR mediated by donor DNA resulted in only 0.5% efficiency with a much higher percentage of

the target DNA instead of a DSB. Mutation in either the RuvC or HNH domains of Cas9 generates a nCas9. BEs generally use nCas9-D10A (mutated RuvC).

Off-target editing: any unintended editing that occurs due to nonspecific interaction of nucleases at sites other than the targeted site in DNA or RNA. Recognition of non-canonical PAM and partial homology of the guide RNA sequence with nontarget sequence may lead to off-target editing.

Protospacer adjacent motif (PAM): a short sequence (2–6 bp) of nucleotides situated immediately adjacent to the target DNA sequence and which is essential for target recognition by CRISPR-associated nucleases. For most discovered nucleases, PAM is present downstream from the target sequence, although some (e.g., Cas12a) can have an upstream PAM.

RNA base editor (RBE): developed by combining catalytically inactive nuclease (dCas13b) and 'adenosine deaminase acting on RNA', RBEs convert a targeted A base to an I base in RNA.

Table 1. Base Editing in Various Organisms[a,d]

Organism/cell type	Target genes	Base editor	Delivery	Editing frequency	Indel frequency	Off-target editing frequency	Refs
Mouse Astrocyte	*Apoe4*	BE3	Plasmid nucleofection	58–75%	4–6%	<0.1%	[9]
Human breast cancer cell line (HCC1954)	*TP53*		Plasmid nucleofection	3–8%	<0.7%	<0.1%	
Human embryonic kidney (HEK)293T cells	Six loci (including *EMX1*, *FANCF*, and *RNF2*)		Plasmid lipofection	5–37%	1–6%	0–19%	
Chinese hamster ovary (CHO) cells	*Hprt*, *Emx1*, *Efemp1*, *Mgat1*	Target-AID	Plasmid lipofection	17–55%	4–12%	<1.5%	[10]
Saccharomyces cerevisiae BY4741	*Can1*		Plasmid transformation	96%	0.2%	<0.1%	
Escherichia coli	*GFP*	ZF-AID	Plasmid transformation	14.4–24.2%	2.77–4.44%	0%	[38]
HEK293FT cells	*GFP*		Plasmid transfection	2.5%	NR	NR	
HEK293T cells	*HEK293 site 3*, *EMX1*, *FANCF*, *VEGFA*	BE3	Plasmid transfection	29 ± 5%	<5%	1.1%	[16]
HEK293T cells	*HEK293 site 3*, *EMX1*, *FANCF*, *VEGFA*	HF-BE3	Plasmid transfection	21 ± 3%	<5%	0.03%	
HEK293T cells	*HEK293 site 3*, *EMX1*, *FANCF/ VEGFA*	BE3	RNP lipofection	26 ± 5%	<5%	<0.025%	
HEK293T cells	*HEK293 site 3*, *EMX1*, *FANCF/ VEGFA*	HF-BE3	RNP lipofection	13 ± 3%	<5%	<0.025%	
Zebrafish embryo	*TYR1* *TYR2* *TYR3*	BE3	RNP lipid nanoparticle injection	5.3 ± 1.8% 4.3 ± 2.1% 0.15–0.54%	0.2–1.9%	NR	
Mouse pups	*VEGFA site 2*	BE3	RNP lipid nanoparticle injection	<1.5%	<0.1%	NS	
HEK293T	*EMX1*, *FANCF*, *HEK2*, *HEK3*, *HEK4*, *RNF2*	Target-AID	Plasmid transfection	6–55%	0.5–3%	NR	[12]
		BE4		18–55%	0.5–6%	NR	
		BE4-Gam		17–58%	<1.5%	NR	
	FANCF, *HEK3-1*, *HEK3-2*, *HEK4*	Sa-BE4		25–60%	<1%	NR	
		Sa-BE4-Gam		42–67%	0.5–4%	NR	
HEK293T	*FANCP*, *HEK293-3*, *HEK293-4*	Sa-BE3	Plasmid transfection	~50–75%	NR	0–35% (FANCF)	[31]
HEK293T	*EMX1*, *FANCF*, *HEK293-3*, *HEK293-4*	SaKKH-BE3		14–62%			
HEK293T	*EMX1*, *FANCF*, *RUNX1*, *HEK293-3*	VQR-BE3		14.5–52%		<1%	
HEK293T		EQR-BE3		7.5–35%		<1%	

Table 1. (continued)

Organism/cell type	Target genes	Base editor	Delivery	Editing frequency	Indel frequency	Off-target editing frequency	Refs
HEK293T	*FANCF*, *HEK293-3*, *HEK293-4*	VRER-BE3		11–32%		NR	
HEK293FT	*Cluc*, *PPIB*, *AVPR2*, *FANCC*, *KRAS*	REPAIRv1	Plasmid transfection	14–38%	NA	1829 off-targets	[86]
	KRAS, *PPIB*	REPAIRv2		13–27%	NA	20 off-targets	
Mouse embryo	*Dmd*	BE3	mRNA injection	44–67%	7.14%	0%	[49]
	Tyr			100%	0%	0%	
	Dmd		RNP electroporation	56%	0%	0%	
	Tyr			57–69%	0%	0%	
Mouse embryo	*Tyr*	HF2-BE2	mRNA injection	11.6–50%	4.54%	0%	[48][c]
Xenopus laevis embryo	*Tyra*, *Tyrb*, *p53*	BE3	RNP injection	5–20.5%	14.6–35.9%	0%	[87]
Zebrafish embryo	*twist2*, *gdf6*, *and ntl*	BE3	mRNA injection	5–20%	NR	NR	[88]
	twist2, *tial1*, and *urod*	VQR-BE3		5–30%	5–35%	NR	
HEK293T	*SPRTN*, *FANCM*, *CHEK2*, *TIMELESS*, and *SMARCAL1*	BE3	Plasmid transfection	21–39%	9%	NR	[60]
HEK293T	*EHMT2* (*G9a*), *LMNB2*	BE3	Plasmid transfection	30–75%	NR	NR	[59]
HCT116 with knocked in mClover	*LMNB2–mClover*			~10.2%	NR	NR	
hIPSC	*PCSK9*	BE3	Plasmid transfection	NR	NR	NR	[54]
Mouse liver	*Pcsk9*	BE3	Adenoviral transformation	19.3–33.6%	0.96–1.96%	0.01–0.2%	
Triponuclear human embryo	*HEK293 site 4*, *RNF2*	BE3	mRNA microinjection	87.5–100%	0.01–0.1%	0–5.35%	[12]
HEK293T	17 Genomic sites	ABE7.8/7.9/7.10	Plasmid transfection	7–70%	<0.1%	0.2–1.3%	[11]
U2OS	Six genomic sites	ABE7.8/7.9/7.10	Nucleofection	8.5–53%	<0.1%	NR	
HEK293T	*HBG1* and *HBG2*	ABE7.10	Plasmid transfection	29–30%	1.2–1.4%	NR	
LCL	*HFE*	ABE7.10	Electroporation	28%	NS	NR	
Fibroblast	*MPDU1*, *SCN9A*	BE4	Plasmid transfection and nucleofection	14–34%	<3.65%	NR	[51]
		BE4max		69–77%	<2.87%		
N2a		AncBE4max		75–84%	<3.73%		
HEK293T	*HBG1/HBG2*	ABE7.10		9–16%	≤1.6%	NR	
		ABEmax		27–52%			

Table 1. (continued)

Organism/cell type	Target genes	Base editor	Delivery	Editing frequency	Indel frequency	Off-target editing frequency	Refs
HEK293T	*HBB*	eA3A-BE3	Lentiviral transfection	22.5%	NR	<0.3%	[42]
		eA3A-HF1-BE3-2xUGI		17.5%			
		eA3A-Hypa-BE3-2xUGI		14%			
Human erythroid precursor cells	*HBB*	eA3A-BE3	RNP electroporation	14.5%	NR	<3.5%	
		A3A (N57Q)-BE3		31.5%		<1%	
HEK293FT	*DNMT3B*, *EMX1*, *PPEF1*, *FAP*, *IGF1*, *MYOD1*, and *IDO1*	BE-PLUS	Plasmid transfection	2–38%	0.3–2.11%	~0.73%	[39]
HEK293T	20 Genomic loci	xCas9-BE3	Plasmid transfection	37% (NGG PAM) 6–24% (NGH) 12% (GAT)	NR	NR	[33]
	Seven genomic loci	xCas9-ABE7.10		69% (NGG PAM) 21–43% (NGM) 16% (GAT)	NR	NR	
HEK293FT	12 Target loci including *FANCP*, *EMX1*, *VEGFA*, *RUNX1*, *DNMT1*	Cas12a-BE	Plasmid transfection	3–46%	0.07–25%	0.009–15%	[15]
U2OS				10–33%	0.11–32%	NR	
Mouse embryo	*Ar*, *Hoxd13*	ABE7.10	mRNA injection	81–100%	0%	0%	[23][b]
	Tyr	SaBE3		69%	27%	0%	
Rabbit blastocyst	*Mstn*, *Dmd*, *Tia1*, *Tyr*, *Lmna*	BE3	mRNA injection	53–88%	0–20%	6–88%	[64]
	Dmd, *Tia1*	BE4-Gam		75–80%	4.3%	4–21%	
	Five loci in *Dmd*, *Otc*, and *Sod1*	ABE7.10		44–100%	0%	0–22%	
Escherichia coli	*tetA*, *GFP*, *rppH*	BE3	Heat shock plasmid transformation	99–100%	NR	NR	[89]
Brucella melitensis	*virB10*		Electroporation	100%	NR	NR	
Bombyx mori	*Blos2, Yellow-e, mCherry, EGFP, puromycin, GAPDH, V-ATPase B*	BE3	Plasmid transfection	3.4–66%	0.6%	NR	[90]
Mouse embryo	*Tyr*	ABE7.10	mRNA injection	13–68%	0%	0%	[24]
Adult mouse	*Dmd*		Intramuscular injection of dual *trans*-splicing viral vectors	3.3	0%	0%	
HEK293T	*Pah*	d*Lb*RR-minBE	Plasmid transfection	4.2%	0–3%	NR	[37]
		dLbRR-BE		23.8%			

Table 1. (continued)

Organism/cell type	Target genes	Base editor	Delivery	Editing frequency	Indel frequency	Off-target editing frequency	Refs
		SaKKH-BE3		46%			
Adult mouse liver		SaKKH-BE3	Dual AAV injection into tail vein	6.1–29.1%	1.1–13.2%	NS	
Murine liver (*in utero*)	*Pcsk9*	BE3	Adenoviral delivery through vitelline vein injection	0.11–15%	2.20%	NS	[74]
	Hpd			0.03–35.84%	4.14%	NS	
HEK293T	20 Target loci	Target-AID-NG	Plasmid transfection	1–38%	NR	NR	[34]
HEK293T, HCT116, HepG2, and MCF7	*RELA*, *PIK3CA*	BE3	Plasmid transfection	6.3–26.4%	NR	<0.1–3.3%	[61]
	BRCA2, *RELA*	VQR-BE3, SaKKH-BE3		0.93–46.6%	NR	0.1–2.9%	
	18 Loci	BE3, VQR-BE3, SaKKH-BE3		0.1–49%	NR	<0.1–33%	
Rat, fertilized egg	*Hemgn*, Two loci of *Ndst4*	ABE-7.10	mRNA injection	90–93%	NS	0%	[91]
HEK293T	β-Catenin	BE3	Plasmid transfection	31%	2%	0 to <0.05%	[55]
			RNP delivery	13%	0.52%		
Postmitotic cells of mouse inner ear			Intracochlear injection of RNP	0.7–2.8%	0.4%	NR	
HEK293T	Four loci	ScCas9-BE3	Plasmid transfection	19–41%	NR	NR	[35]
	One locus	ScCAs9-ABE7.10		21%	NR	NR	
HEK293T	*DMD*, *AAVS*, *HBB02*, *FANCF02*, *EMX1*	Sniper-Cas9-BE3	Plasmid transfection	0.04–50%	NR	<5% for EMX1	[92]
Mouse embryo	*Hbb-bs*, *Fah*, *Gaa*	ABE7.10	mRNA injection	52–100%	NR	<0.2%	[93]
	Hbb-bs	VQR-ABE		20%	NR		
Rat embryo	*Otc*	Sa-KKH-ABE		16%	NR		
HEK293T	15 Methylated (native high and low) loci from *FANCF*, *MAGEA-1*, *MSSK-1*, *PDL1*, *VEGFA*	hA3A-BE3, hA3A-BE3-Y130F, hA3A-BE3-Y132D	Plasmid transfection	2–62%	0.5–10%	NR	[41]
NIH/3T3	Six loci from *APC*, *PIK3CA*, *CR8.OS2*	BE3 (RA)	Lentiviral transduction	30–58%	0.4–3%	NR	[46]
		FNLS-BE3		41–93%	0.1–0.63%	NR	
Mouse intestinal organoids	*APC*, *PIK3CA*	FNLS-BE3	Plasmid transfection	47.5–97.2%	<1%	NR	
Adult mouse liver	*Ctnnb1*	FNLS-BE3	Hydrodynamic delivery of plasmid	NP	NR	NR	
Zebrafish embryo	*chd*, *oep*	Target-AID	mRNA injection	2.19– 4.37%	8.48%	20%	[47]

Table 1. (continued)

Organism/cell type	Target genes	Base editor	Delivery	Editing frequency	Indel frequency	Off-target editing frequency	Refs
HEK293T	*FBN1*	BE3	Plasmid transfection	40%	NR	0%	[73]
Human embryo			mRNA injection	80–100%	0%		
Zebrafish embryo	*atp5b, rps14, wu:fc01d11, musk*	ABE7.10	mRNA injection	8.30–22.22%	7–24%	0%	[94]
				ABE7.10max	19.2– 40.7%		
3.4–4%	13 Loci including *Wap*, *Csn2*, *Tyr*, and *Dmd*						
Mouse zygotes		HF2-BE2	mRNA injection	44.68%	4%	36%	[25]
		BE3		70.32%	12%	0%	
		SaBE3		58.8%	29%	6%	
		VQR-BE3		91.3%	4%	22%	
		BE4		85%	19%	4%	
		ABE		100%	0%	0%	
HEK293T	*VEGFA*	Spy-mac-BE3	Plasmid transfection	20–30%	NR	NR	[36]
Arabidopsis	*ALS*	BE3	*Agrobacterium*-mediated floral dip	1.7%	NR	NR	[95]
Rice	*PDS*, two loci in *OsSBEIIb*	BE3	*Agrobacterium*-mediated	0.1–20%	0–9.61%	NR	[96]
Rice	*NRT1.1B, SLR1*	BE3 (-UGI)	*Agrobacterium*-mediated	0–13.3%	~10%	NR	[44]
Rice	*CDC48*	BE3	*Agrobacterium*-mediated	43.5%	0%	0%	[97]
Maize	*CENH3*			10%	NR	NR	
Wheat	LOX2		Particle bombardment	1.25%	0%	0%	
Rice	*ALS*, *FTIP1e*	Target-AID	*Agrobacterium*-mediated	6–89%	10–62%	0%	[76]
Tomato	*DELLA*, *ETR1*			41–92%	16–69%	0.14–0.38%	
Rice	*CERK1*, *SERK1*, *SERK2*, *ipa1*, *Pi-ta*, *BRI-1*	BE3	*Agrobacterium*-mediated	0–38.9%	0%	NR	[45]
Rice	*ACC*, *ALS*, *CDC48*, *DEP1*, *NRT1.1B*	ABE7.10	*Agrobacterium*-mediated	3.2–59.1%	0%	0% for ACC	[27]
Wheat	*DEP1, GW2*		Particle bombardment	0.4–1.1%	0%	NR	
Rice calli	*MPK6*, *MPK13*, *SERK2*, *WRKY45*, *Tms9-1*	ABE7.10 ABE7.8	*Agrobacterium*-mediated	0–62.26%	0%	0%	[28]
Arabidopsis	*FT*, *PDS3*	ABE7.10	*Agrobacterium*-mediated	0–85%	NR	<0.4%	[98]
Brassica napus	*ALS*, *PDS*	ABE7.10, ABE6.3, ABE7.8, ABE7.9	Protoplast transformation	8.8%	<0.1%	NR	
Rice		ABE7.10		12.5–26%	0%	0%	[26]

Table 1. (continued)

Organism/cell type	Target genes	Base editor	Delivery	Editing frequency	Indel frequency	Off-target editing frequency	Refs
	SPL14, *SLR1*, *SPL16*, *SPL18*		*Agrobacterium*-mediated				
	SPL14, *SPL17*, *SPL16*, *SPL18*	ABE-Sa (SaCas9-nickase)		17–61%	0%	0%	
Rice	*SPL13*, *SPL14*, *SPL16*, *SPL17*, *SPL18*, *GRF4*, *TOE1*, *IDS1*, *MTN1*, *SNB*	ABE-Sa ABE-VQR ABE-VRER ABE-SaKKH	*Agrobacterium*-mediated and particle bombardment	0–74.3%	NR	NR	[99]
	PMS1, *PMS3*, *SNB*	BE3 VQR-BE3 SaKKH-BE3		0–80%	0%	NR	
Wheat	*ALS*, *MTL*	hA3A-BE	Particle bombardment	16.7–22.5%	0%	0%	[40]
Rice	*CDC48*, *NRT1.1B*		*Agrobacterium*-mediated	44–83%	0%	0%	
Potato	*GBSS*		Protoplast transformation	6.5%	0%	0%	
Rice Calli	*EPSPS*, *ALS*, *DL*	Target-AID-NG	*Agrobacterium*-mediated	5–95.5%	0–68%	NR	[100]

[a]For calculating the editing frequency, only anticipated products are considered. Protoplast assay (in plants) and preliminary studies are not included.
[b]Proximal off-targets are mentioned.
[c]Indel frequencies were determined for other loci (not mentioned in the table).
[d]Abbreviations: BE2, rAPOBEC1-dCas9-UGI; BE3, rAPOBEC1-nCas9-UGI; BE4, rAPOBEC1-nCas9-UGI-UGI; HEK293T, human embryonic kidney cells 293T; HF2, high-fidelity version of dCas9; hIPSCs, human induced pluripotent stem cells; Lb, *Lachnospiraceae bacterium;* LCL, lymphoblastoid cell line; N2a, neuroblastoma cells; NA, not applicable; NIH/3T3, mouse fibroblast cell line; NP, near-perfect; NR, not reported; Nucleotide M, A/C; Nucleotide H, A/C/T; RA, codon-optimized; FNLS (RA + extra NLS + FLAG tag); Sa, *Staphylococcus aureus*; Sc, *Streptococcus canis*; Sp, *Streptococcus pyogenes*; Spy-mac, Cas9 fusion derived from *Streptococcus pyogenes* and *Streptococcus macacae*; U2OS, human osteosarcoma cells.

indels (4.3%) [9]. So far, few studies have reported a significant amount of indel formation with CBE (Table 1). During cytosine base editing, incorporation of U in DNA is treated as an error/damage by the cellular base-excision repair (BER) machinery, and is removed by the ubiquitous uracil-DNA glycosylase (UDG) [20]. Removal of a U by UDG causes the formation of an abasic site and subsequent repair by error-prone polymerases leading to random nucleotide incorporation. In addition, occasional strand breaks might be responsible for indel generation [21]. Indel formation was decreased 7–100-fold in UDG knockout cell lines [12], and coexpression of UDG inhibitor (UGI) with CBE improved base-editing efficiency and reduced indel formation [9,10].

In comparison with CBEs, ABEs yield a much cleaner edited product that has virtually no indels (Table 1). Although alkyl adenine DNA glycosylase is known to counteract the incorporation of inosine in DNA [22], inosine excision was not found to impede ABE performance [11]. This might be one of the strong reasons behind the superior performance of ABEs than CBEs in terms of product purity and editing efficiency. ABE recovered $\geq$99.9% pure product with a negligible rate of indels ($\leq$0.1%) [11]. Whole-genome sequencing revealed no off-target base editing by ABE in studies on mouse models [23,24]. ABEs also exhibit a remarkably higher fidelity to generate the anticipated base editing in mouse embryos than do CBE variants [25].

Similarly, ABE edited rice and wheat plants were recovered without any undesired edits at on- and off-target sites of the genomes [26–28]. Surprisingly, recent whole-genome sequencing studies in mouse and rice detected that BE3 could induce a significant amount of off-target C → T editing, mostly in the transcribed region, even in absence of guide RNA [29,30]. The results indicate that the cellular presence of cytosine deaminases and their encounter with unwound single-stranded (ss)DNA may cause untargeted C → T conversion in the genome. Both studies also reported that ABE-induced off-target editing events are rare throughout the whole genome, at a level that is not significantly higher than the rate of spontaneous mutation [29,30]. Nevertheless, the off-targeting outcomes for BEs and Cas9 alone are not always the same, which raises the need for separate evaluation criteria and methods. A modified di-genome sequencing and targeted deep sequencing method was found to be highly sensitive to assess the specificity of CBE [17]. A new study also reported a novel and precise method of off-target detection 'genome-wide off-target analysis by two-cell embryo injection' (GOTI) followed by whole-genome sequencing [29].

BEs with Alternative PAM Compatibility

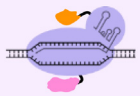
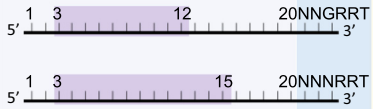
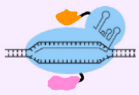
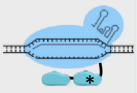
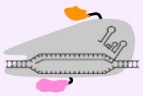
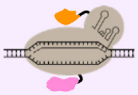
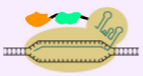
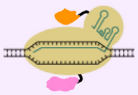
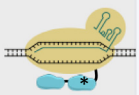
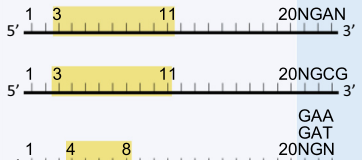
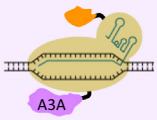
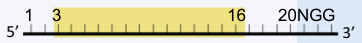
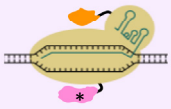
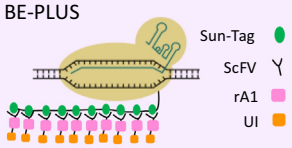
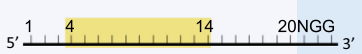
A key limitation to using BEs is the requirement of a suitably positioned NGG **protospacer adjacent motif** (PAM) that keeps the target C/A within a narrow (~5 base) **activity window**, which dramatically restricts the number of targetable sites [31]. To address this limitation, several groups of researchers have developed improved versions of BEs which have either alternative PAM requirements or relaxed PAM compatibilities (Figure 1). Kim and colleagues developed a set of CBEs by replacing the SpCas9 with other Cas9 variants that permit targeting of genomic loci with suitably placed NGAN, NGCG, or NNNRRT sequences [31]. BEs with alternative PAMs have also been developed for plant systems [32].

Recently, Hu and colleagues evolved SpCas9 to generate a variant, xCas9, which has PAM flexibility that can accept NG, GAA, or GAT sequence [33]. xCas9-based CBEs and ABE developed in the study expanded the scope of base editing at genomic sites that were previously inaccessible [33]. Similarly, an engineered version of Cas9 has been developed, SpCas9-NG, that can accept all NG PAMs [34]. The BE (nCas9-NG-AID) developed using the SpCas9-NG showed substantial C → T conversion efficiency at all NG PAM sites tested. Recently, ScCas9- and Sp-macCas9-based BEs have also been created to target 5′-NNG-3′ and 5′-NAA-3′ PAM sequences, respectively [35,36]. However, the use of BEs based on xCas9, SpCas9-NG, and ScCas9 would increase potential off-target editing owing to their broader PAM compatibility.

SpCas9-based BEs are limited to targeting mainly GC-rich genomic regions. This short-coming was addressed recently by developing CBE based on Cas12a which requires a T-rich PAM (5′-TTTV-3′) [15,37]. Interestingly, zinc finger (ZF)- or transcription activator-like effectors (TALE)-based CBEs have also been developed by tethering the DNA-binding domains of the proteins with either APOBEC1 or AID [38]. These versions can be engineered to target any desired genomic sequences because ZFs and TALEs do not require a PAM.

BEs with Shortened or Extended Activity Windows

Using the conventional BEs (Target-AID, BE3, and ABEs) (Boxes 1 and 2), one can target bases (C/A) present at protospacer positions 2–6 or 4–8 bp distal to the PAM [9–11]. Figure 1 represents BEs with diversified editing windows. Recently, a CBE platform, BE-PLUS, with a SunTag signal amplification system, displayed an expanded editing window in the range of 1–14 bp [39]. The BE-PLUS can theoretically recruit 10 copies of APOBEC1-UGI per

CBE	ABE	Windows	PAM	Nuclease variants	Refs
			20NNGRRT 20NNNRRT	SaCas9(n) SaCas9-KKH(n)	[23,26, 31,32, 37,61]
			20NNG	ScCas9(n)	[35]
			1 VTTT	LbCas12a(d)	[15,37]
			20NAAN	Sp-macCas9(n)	[36]
			20NGN	SpCas9-NG(n)	[34, 100]
			20NGAN 20NGCG GAA GAT 20NGN	SpCas9-VQR(n) SpCas9-VRER(n) Sp-xCas9(n)	[31–33, 61,88]
A3A-BE3 			20NGG	SpCas9(n) Sp-xCas9(n)	[40–42]
YE1-/YE2-/EE-/YEE-BE3 			20NGG	SpCas9(n)	[31]
BE-PLUS 			20NGG	SpCas9(n)	[39]

Trends in Biotechnology

Figure 1. Base Editors with Diverse Protospacer Adjacent Motif (PAM) Compatibilities and Activity Windows. The column highlighted in pink encompasses all cytosine base editors (CBEs), while grey highlights adenine base editors (ABEs). The CBE platforms use a Cas D10A nickase (n) (variants are colored distinctly) or deactivated Cas (d) (grey) tethered with cytidine deaminase [rAPOBEC1 (rA1) (pink), or PmCDA1 (lime green), or hAPOBEC3A (A3A) (violet)], and

(Figure legend continued on the bottom of the next page.)

nCas9. Hypothetically, the extended base-editing window of BE-PLUS is due to exposure of the entire length of the non-complementary ssDNA of the R-loop to these 10 copies of deaminases. The base-editing window of Cas12a-CBE ranges from 8 to 13 in protospacer position when the base next to PAM is counted as 1 (unlike Cas9, Cas12a has a 5′ PAM) [15]. The editing window of Cas12a-CBE was reduced to 3–4 bp by using a mutated deaminase [37]. Human APOBEC3A (hA3A)-based CBE exhibited a much wider editing window, from protospacer positions 1 to 17, in rice, wheat, and potato plants [40]. Another study showed that hA3A-CBE has a ~12 bp editing window [41]. However, Gehrke and colleagues reported that the editing window of hA3A-CBE was only 5 bases in length in a mammalian system [42]. Although these studies used the same XTEN linker between nCas9 and hA3A, it remains unclear why the editing window length is almost threefold longer in the plants than in the mammals. Clearly, the editing window boundaries of BEs can vary from target to target [11]. Nevertheless, these new additions to the base-editing toolkit would enable targeting Cs more proximal to the PAM.

Another practical difficulty in using BEs is when additional adjacent C/A bases are present with the target base in the 5–6 bp activity window because simultaneous editing of those bases (bystander editing) would result in undesired mutations. Use of mutant APOBEC1 or truncated CDA1 led to the development of CBE variants with a narrowed editing window [31,43]. Although the variants in these studies shortened the editing window from ~5 to ~2–3 bp, the complication of editing multiple Cs in the narrowed window remains. The issue has recently been addressed by replacing the rAPOBEC1 enzyme of BE3 with an engineered hA3A (eA3A) [42]. The eA3A-BE3 prefers the hierarchy of TCR > TCY > VCN motifs for deamination, and greatly reduces the editing efficiency at other Cs in the editing window, although it exhibits similar performance to BE3 on C in TC motifs. It is also noteworthy that eA3A was found to be equally efficient when fused with Cas9 variants (xCas9 and VRQR-Cas9) with broader PAM compatibility [42]. This development significantly increases the scope of genome-wide applicability.

Scientists have attempted to manipulate the length of the editing window by varying the protospacer length (truncated or extended), but without success [10,27,31]. However, the window was found to be slightly broadened with an extended protospacer [24]. In addition, variation in the length of the linker (usually 32 amino acids) between the deaminase and nCas9 does not alter the width of the editing window [31,42]. Nevertheless, use of a stringent proline-rich linker has been reported to narrow the window [43].

Sequence Preferences

BE3 prefers some sequences over others for its deamination activity. If a target C is present immediately downstream of a G, BE3 editing was found to be inefficient [9,15]. BE3 displays sequence preferences in the order TC ≥ CC ≥ AC > GC. However, CDA1-BE3 (PmCDA1) and AID-BE3 (hAID) more efficiently edit those GC genomic loci than does BE3, albeit the two versions displayed lower efficiencies at other genomic loci [12]. However, hA3A-BE3 has been demonstrated to display efficient editing independently of sequence context [40].

uracil glycosylase inhibitor [UGI/UI (orange)]. The ABE platforms use nCas9(D10A) with laboratory-evolved deoxyadenosine deaminase [heterodimeric ecTadA (teal)] [11]; *, denotes mutated. The estimated activity windows are shown in the same color codes as the nucleases. The first nucleotide of PAM is numbered as 21 (except for Cas12a-BE3). Nucleotides: N, A/T/G/C; R, A/G; V, A/G/C. Abbreviations: Lb, *Lachnospiraceaebacterium*; Sa,*Staphylococcus aureus*; Sc, *Streptococcus canis*; ScFv, single-chain variable fragment; Sp, *Streptococcus pyogenes*; Sp-macCas9, hybrid Cas9 of *S. pyogenes* and *Streptococcus macacae*.

An engineered hA3A (eA3A)-based CBE has been developed to reduce editing in motifs other than TC [42].

Owing to the native preference of *Escherichia coli* TadA, early versions of ABEs (ABE1s–ABE5s) had a sequence preference for YAC (Y = C/T) and displayed reduced editing efficiencies at multiple A-containing target loci. However, new versions of ABEs (ABE6.3, ABE7.8, ABE7.9, and ABE7.10) were evolved to overcome sequence limitations and provide broader sequence compatibility [11].

Site-Specific Ineptness

Some genomic sites may not be accessible to the BEs, probably owing to preoccupancy with other proteins or nucleosome. Recently, a site in rice *PMS1* has been shown to be resistant to BEs, whereas the same site was accessible to wild-type Cas9 [99]. In our study, a target site in rice *Wsl5* gene had a very low-efficiency of A → G base editing compared with another genomic site (K.A.M. and Y.Y., unpublished data). Some single-guide (sg)RNAs with BEs may not be as active as they are with active Cas9 alone, or vice versa [17]. Canonical BE3 showed decreased efficiency of editing when the target C is embedded in a highly methylated region, whereas hA3A-BE3 exhibited ~threefold improvement in performance over BE3 at those sites [41]. This development increases the likelihood of successful editing for the ~43% of disease-related C → T mutations which lie in the context of CpG islands [41].

Improved BEs To Reduce Undesired Point Mutation

CBE at some genetic loci are reported to generate undesired point mutations (C → A or C → G) other than the targeted C → T conversion [25,31,44–47]. In addition to its expanded editing window, the recently developed BE-PLUS offers more high-fidelity base editing than BE3 by inducing fewer indels and undesirable substitutions (C → A/C → G) [39]. Two high-fidelity CBEs (eA3A-HF and eA3A-Hypa) were demonstrated to reduce unwanted C → G mutations [42].

Off-target editing by BEs occurs mainly due to nonspecific interactions of nuclease with the genome. High-fidelity Cas9 (HF-Cas9), a mutated Cas9 variant, was employed to generate HF-BE3 to reduce off-target editing [16]. Although HF-BE3 reduced the mean off-target editing by 37-fold relative to BE3, it exhibited a slight reduction in on-target base-editing efficiency. Surprisingly, HF2-BE2 generated cytidine substitutions even 3 bp downstream or 38 bp upstream of the gRNA target site [48]. Likewise, a recent report of substantial genome-wide off-target mutation by HF1-BE3 with or without gRNA [30] raises the question: how much does the nuclease domain of a CBE contribute to off-target modification? Fourth-generation CBE (BE4), constructed by including an extra copy of UGI and extending the N- and C-terminal linker of nCas9 in the BE3 background, displayed substantial improvement in editing efficiency and product purities [12]. Reduced indel formation by BE4 further supports the view that a cellular 'uracil BER mechanism' plays a significant role in indel formation. Fusion of bacteriophage Gam protein, a nuclease inhibitor that stabilizes DSB ends, with BE3 and BE4 further limits indel formation [12]. Coexpression of free UGI with BE3 and fusion of three copies of UGI in the BE3 architecture dramatically reduced the generation of C → A/C → G substitutions and indels relative to the original BE3 [50].

Expression Matters

The base-editing efficiency at specific target genomic loci or in some cell types is sometimes limited because of low expression and the inadequate availability of BEs [51]. Early CBEs (BE3

and BE4) and ABE proteins were targeted to the nucleus using the SV40 nuclear localization signal (SV40-NLS) fused at the C terminus [9,11]. However, a bipartite SV40-NLS has been shown to outperform SV40-NLS regarding Cas9 nuclear targeting and subsequent editing [52]. Addition of a bipartite NLS to both the N and C termini of BE4 and the use of a different codon-optimization method generated BE4max which shows significantly higher expression and improved editing efficiency over BE4 [51]. Similar improvement on ABE7.10 has resulted in a superior base-editor, ABEmax [51]. The study also dealt with ancestral sequence reconstruction of APOBEC protein for improving expression while keeping the catalytic activity unharmed. AncBE4max, developed by replacing the rAPOBEC1 in BE4max with ancestral APOBEC1 (Anc689), exhibited further enhancement in editing performance. Similarly, codon optimization for mammalian expression and the addition of an extra NLS resulted in up to 30–50-fold increased editing over BE3 [46]. The level of expression might also play a significant role in editing outcomes such as in monoallelic versus biallelic point mutation.

An exogenously controlled system of base editing has been developed using 'aptazyme-embedded' guide RNAs [53]. A small complementary sequence blocked the guide RNAs in the absence of ligand theophylline. The presence of theophylline induces aptazyme-mediated cleavage of the blocking sequence and subsequently activates the guide RNA and base editing. This inducible base-editing system could be applied for specific genome engineering purposes.

Broad Applications in Basic Biology, Medicine, and Agriculture

Precise Editing in Non-dividing Cells

Precise editing in terminally differentiated cells is not readily achievable through HDR because HDR is restricted to the S and G2 phases of the cell cycle. Given that base editing is dependent on cellular mismatch repair machinery rather than on recombination mechanism [9–11], it offers a potential alternative method to create point mutation in non-dividing cells. *In vivo* delivery of BEs yielded successful targeted mutations in adult mouse liver cells [37,46,54]. Yeh and colleagues demonstrated base editing in postmitotic mouse inner-ear cells by installing a β-catenin gene mutation to upregulate Wnt signaling [55]. Similarly, ABE generated targeted A → G editing in an adult mouse model [24]. Base editing can be applied to reverse a percentage of genetic defects in non-dividing cells affected by ~5000 known human monogenic disorders [56].

Directed Gene Evolution and Genomic Diversification

In the absence of UGI, cytidine deaminase can also produce mutations other than C → T [12,57,58]. Diverse libraries of targeted point mutations could therefore be generated using CBE lacking a UGI and then screened for desired gene function. Two independent studies (CRISPR-X and TAM) utilized this strategy by fusing hAID with dCas9 to generate large pools of functional variants [57,58]. They were able to identify novel mutations in the *BCR–ABL* and *PSMB5* genes that confer resistance to chemotherapeutic drugs. The identification of gene sequence variants resistant to drugs will expedite future drug design and development. These two proof-of-concept studies demonstrate that base editing can be utilized for the directed evolution of biomolecules and for generating new libraries with diverse applications in industrial engineering, synthetic biology, and many other fields [7] (Figure 2A, Key Figure). Although active Cas9-mediated library development could generate variants, the lion's share of the mutants may contain indels and frameshifts which are sometimes undesirable.

Creation and Correction of Early Stop Codons

In DNA, there are three stop codons, ochre (TAA), amber (TAG), and opal (TGA) in the standard genetic code. Because CBE converts C → T (G → A in the opposite strand), it can switch

Key Figure

Broad Application and Utility of Base-Editing Technologies

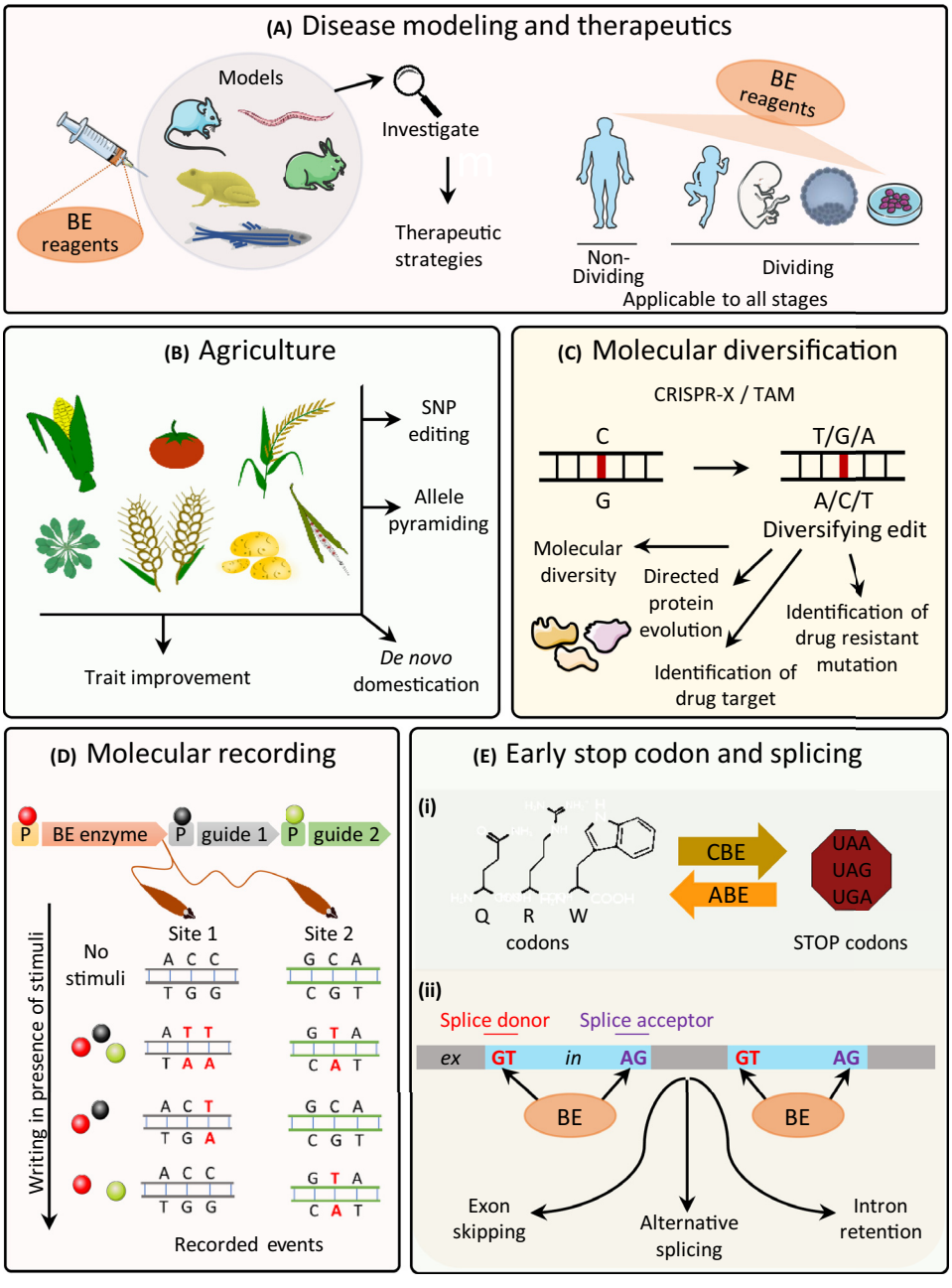


Figure 2. (A) Disease modeling and therapeutics: Base editor (BE) tools have been successfully used to create animal models of human disease. Disease models can be used as an assay system for drug screening and therapeutic strategy development. BE-mediated gene correction has been demonstrated in almost all stages of development including *in vitro*

(Figure legend continued on the bottom of the next page.)

glutamine (CAA, CAG), arginine (CGA), and tryptophan (TGG) coding sequences to stop codons, thus causing premature termination of translation (Figure 2Ei). Two recent studies (CRISPR-STOP and iSTOP) made use of BE3 to knock out (KO) gene function by creating early stop codons [59,60]. According to the predictions of these two studies, 17 000 human genes could be knocked out by introducing early stop codons, and 97–99% genes of seven other eukaryotic model species can be targeted. For gene KO studies, CBE represents an effective and more precise alternative to the widely used functional Cas9, which creates DSBs that results in unpredictable repair outcomes.

On the other hand, ABE can be utilized to correct a defective allele containing a mutation that produces an early stop-codon. All three stop codons could be corrected to tryptophan (TGG) by targeting the coding DNA strand, and to glutamine (CAA, CAG) and arginine (CGA) by targeting the noncoding strand (Figure 2Ei).

Identification of Conserved Functional Amino Acids

By employing both types of BEs, almost all codons can be targeted to create either missense or nonsense mutations, and this permits the identification of conserved amino acid(s) that are crucial for the function of a protein, and can also rapidly validate a large number of algorithm-predicted conserved amino acid residues of biological significance *in vivo*.

Creation of Splice-Site Variants

One of the crucial post-transcriptional gene regulatory processes is alternative splicing. Canonical eukaryotic introns junctions have 5′-GT (the donor sequence) and AG-3′ (the acceptor sequence) which play a fundamental role in splicing. Theoretically, both the highly conserved sites of an intron can be mutated using either of the BEs (CBE/ABE) if a PAM is located at a suitable distance (Figure 2Eii). By employing CBE and ABE, donor/acceptor sites can be disrupted to AT/AA and GC/GG, respectively. Recently, Gapinske and colleagues demonstrated artificial exon skipping through mutating intron acceptor sites using CBE [61]. They estimated that their CRISPR-SKIP approach could target ~63% of internal exons in protein-coding transcripts. Many important human diseases are caused by defects in splicing, and exon skipping is therefore potentially applicable for treating those diseases [61,62]. In addition to exon skipping, BEs can also mediate intron retention and both 5′- and 3′-alternative splice-site generation. Recently, CBE-mediated modulation of RNA splicing and restoration of dystrophin function in mammalian cells have been reported [63,64]. In addition to targeting protein-coding genes, CRISPR-SKIP can be used to study the function and regulation of long noncoding RNAs, where CRISPR-STOP/iSTOP cannot be used [61]. Numerous recent studies revealed that alternative splicing allows plants to rapidly adjust to environmental stress by modulating key elements of the stress-response proteome [65]. Employing CBE, four

cultured cells, zygotes, embryo, *in utero* fetus, newborn pups, adult liver cells, postmitotic cells in the ear, etc. (B) Agriculture: BEs can play a significant role in crop improvement. Base editing has been demonstrated in major agricultural crops such as rice, wheat, maize, tomato, potato, and *Brassica*. (C) Molecular diversification: CRISPR-X, and TAM technologies have been used in artificial evolution and diversification of protein structure and function. Because the two platforms lack UGI, they can convert target C to T/G/A. (D) Molecular recording: BEs can be used to write base substitutions in targeted loci in response to external or internal stimuli and to record stimulus-responsive molecular events. The CAMERA and DOMINO platforms have utilized BEs to record cellular memory. Red nucleotides result from BE activities. (E) Creation and correction of an early stop codon and modification of alternative splicing: the cytosine base editor (CBE) is used to mutate glutamine (Q), arginine (R), and tryptophan (W) codons to stop codons, whereas the adenine base editor (ABE) can be used to reverse the action. BEs can be used to disrupt canonical splice donor (GT) and acceptor (AG) sites in a gene to interfere with normal splicing and the generation of RNA variants. Abbreviations: *ex*, exon; *in*, intron, UGI, uracil-DNA glycosylase (UDG) inhibitor.

Arabidopsis genes have been functionally revalidated by either modifying constitutive splicing or impeding alternative splicing [66]. Hence, splice-site mutation by BEs has broad application not only in gene therapy but also in basic studies across the plant and animal kingdoms.

Gain of Function/Loss of Function at Single-Base Resolution

Study of genetic gain-of-function and loss-of-function mutations enables scientists to unravel the functional details of a wide array of genes, which in turn advances our knowledge in medical sciences and crop improvement. Convenient technologies to study loss-of-function mutations have been available since the discovery of RNAi. CRISPR/Cas9-mediated gene disruption and CRISPR interference (CRISPRi) have greatly expedited loss-of-function studies [19,67,68]. More recently, BE-mediated tools (CRISPR-STOP and iSTOP) to generate early stop codons have further increased the precision of functional analysis. However, for gain-of-function studies, there was no suitable tool until the development of BEs, which can mutate a non-functional SNP to a functional one. Although CRISPR-activator (CRISPRa) can enhance the transcription rate of some genes, it cannot truly help in gain-of-function studies where a gene is non-functional as a result of premature translation termination or that is inactive due to a single or several missense mutations.

BEs as DNA Writers and Molecular Recorders

DNA is superior in many aspects to other media for digital information storage [69]. BEs offer a unique platform to write base substitutions into targeted genetic loci, record them in living cells, and read out using sequencing methods (Figure 2D). CBEs have recently been used to develop CAMERA and DOMINO systems to write biological information onto DNA [70,71]. Both the CAMERA and DOMINO platforms have been demonstrated to successfully record dynamic DNA modification at single-base resolution in response to chemical or physical stimulus as well as their exposure times. BEs provide higher scalability and outcome predictability than other available precise DNA writers [72]. Taking advantage of the reverse directionality of base editing by ABEs over CBEs, the inclusion of ABEs could be utilized to extend the spectrum and complexity of the recording system [14,72].

Therapeutic Applications

BEs have remarkable potential for use in the correction of disease-causing mutations in the human genome. A total of ~20 580 human pathogenic SNPs can be corrected by A–T to G–C or C–G to T–A mutations [14]. Gene therapy is a major area where base-editing reagents could be practically applied because they have already been adopted to study, model, and repair various debilitating human genetic disorders [13]. Mutation in the *Apoe4* gene for Alzheimer's disease in mouse astrocytes and in the *Tp53* gene in a human mammary cancer cell line were corrected using CBE [9]. Additional studies followed this initial example and demonstrated correction of pathogenic mutations in animal models, human cell lines, and even in human zygotes (Table 1). For example, ABE and CBE have been employed to improve muscle function by correcting a premature stop codon and exon-skipping in the *Dmd* gene in a Duchenne muscular dystrophy mouse model, and in patient-derived induced pluripotent stem cells (iPSCs), respectively [24,63]. CBE-mediated correction of a human β-thalassemia promoter mutation and reversion of a human pathogenic mutation causing Marfan syndrome have been reported [42,73]. In addition to correcting disease-causing mutations, BEs could play a significant role in generating different animal models for numerous human diseases, which in turn would greatly facilitate basic study and drug development (Figure 2A). Liu and colleagues generated mouse models by installing A → G pathogenic mutations in the *Ar* and *Hoxd13* genes [23]. More recently, two proof-of-concept studies demonstrated CBE-mediated correction of phenylketonuria and tyrosinemia in adult mouse liver and mouse fetus, respectively [37,74]. Base editing should be possible at almost all developmental stages ranging from oocytes, embryos, and fetuses to adults, indicating the vast potential of therapeutic gene

Box 3. DNA-Free Base Editing To Reduce Regulatory Concerns and Overcome Vector Size Limitation

Delivery of base-editing reagents has relied on genetic transformation of plasmid DNA and its subsequent *in vivo* transcription and translation. However, there is a concern about the uncertain regulatory fate of genome-edited organisms in different countries, especially for crop plants. The Court of Justice of the European Union (ECJ) recently stated that CRISPR-mutagenesis-derived organisms are subject to the same regulations as genetically modified organisms (GMOs). By virtue of genetic segregation over the generation, mutated plants could be obtained with no exogenous DNA fragments (Cas9, sgRNA, and marker gene expression cassettes) used during the genetic transformation. However, the same is more difficult to achieve for asexually reproducing plants such as banana, potato, and many other crop species [85]. Previous studies have documented that delivery of Cas9/sgRNA RNPs into animal and plant cells yielded efficient genome modification [85,105,106]. Delivery of purified BE3 protein complexed with *in vitro* transcribed sgRNA has led to successful targeted base editing with almost unmeasurable off-target editing in zebrafish and mouse models [16]. BE3 RNP was shown to mutate a phosphorylation site of β-catenin in postmitotic cells in mouse inner-ear [55]. Instead of RNPs, injection of CBE mRNA and sgRNA into the cytoplasm of mouse zygotes generated mutant mice with an efficiency as high as ~63%, although with a considerable amount of off-target editing and indel formation [107]. By contrast, in a rat model, *Hemgn* and *Ndst4* gene loci were base-edited at a very high efficiency using ABE mRNA plus sgRNA injection, with minimal off-target editing and indel formation [91]. Higher indel formation and off-target editing in the study with CBE [107] is probably due to the inherent high activity of cellular BER, whereas this was not the case with ABE in the later study [91]. This DNA-free base editing mediated by either RNPs or RNAs has great potential to be a method of choice to avoid the regulatory red tape and to address public concerns.

The DNA-free base editing may also address the issue of the size limitation associated with adeno-associated virus (AAV) vector-mediated base-editor construct delivery for therapeutic application. The cargo size of AAV is limited to <5 kb [108], but all of the base-editing platforms are more than 5 kb in size, and this fact limits the development of human therapeutics. Although this issue has recently been addressed by developing the intein-split BE [37], RNP and RNA delivery provides a straightforward alternative.

editing (Figure 2A). Moreover, demonstration of DNA-free base editing raises great hopes for its therapeutic applications because it may address regulatory concerns (Box 3).

Crop Improvement

Taking rice as a model, 65% of SNPs in coding sequences comprise either C/T or A/G transitions [75], indicating a high likelihood that BEs may be used for crop improvement. Base editing has already been demonstrated successfully in the major crop plants (Table 1). With recent advances in genomics, an increased number of functional SNPs associated with important agronomic traits are being discovered across crop plant species. Some of these SNPs can be installed by BEs in cultivated crop varieties to introduce desired traits (Figure 2B). Generating or pyramiding favorable allelic combinations in a single background is challenging and often takes many years. BEs can do the same job within a much shorter period. Unlike traditional breeding, the use of BEs can also eliminate the problem of linkage drag. Shimatani and colleagues developed herbicide (IMZ)-resistant rice using CBE, while Li and coworkers employed ABE to generate a different herbicide (Haloxyfop-P-methyl)-resistant rice [27,76]. Another vital area where base editing could play a significant role is in breeding for disease resistance. Many plant resistance genes are allelic in nature, differing in single or a few nucleotides. Some alleles act as pseudogenes because of the presence of a nonsense mutation, and if corrected would be able to impart resistance. Likewise, the coding sequences of plant disease susceptibility (*S*) genes [77] can be distorted by incorporating early nonsense mutations using BEs. *S* genes may play various cellular functions, and disruption of an *S* gene may therefore give rise to pleiotropic effects such as changes in growth rates, reduced yield, and sensitivity to other stresses [77]. In the plant–*Xanthomonas* interaction, transcription of some plant *S* genes is enhanced by promoter binding by TAL effectors (TALEs) secreted from bacterial pathogens. The repeat variable di-residue (RVD) of TAL proteins has highly specific binding sequences in the promoter region of *S* genes [78]. Instead of disrupting the coding sequence, the nucleotide/s of the TALE binding site in the *S* gene promoters can be mutated utilizing BEs to enhance resistance without pleiotropic effects.

Unlike the random mutagenesis-mediated TILLING technique, which generates limited mutation density for a gene of interest, the use of BEs could yield a 'high-density mutant population' and facilitate the artificial evolution of agronomically relevant loci [79]. *In vitro* evolution of the NBS/LRR domain using error-prone PCR has been demonstrated to enhance the function of a potato resistance (*R*) gene towards broad-spectrum resistance [80]. Achievement of broad-spectrum resistance has been found to be associated with plant fitness cost, which could be eliminated by further random mutagenesis [81]. The ability to generate almost all missense mutation by BEs qualifies them as an attractive tool to perform targeted evolution of *R* genes while avoiding fitness costs in crop plants.

Concluding Remarks and Future Perspectives

Although base-editing technology has been rapidly advanced by fine-tuning the architecture of BEs to increase the efficiency, targetability, and purity of the edited product, there remain many challenges to be overcome before its full potential can be realized (see Outstanding Questions). Many BEs are available with nickase and dead-nuclease variants that allow more specific genomic editing (Figure 1), but they are not always as efficient as the original BEs developed with SpCas9 (D10A). Further evolution of Cas9 proteins and discovery of new nucleases with more PAM plasticity would broaden the scope of genome targeting while maintaining editing efficiency. For example, for editing T-rich genomic regions, a Cas12a-ABE could be developed but does not yet exist. Safe and effective delivery of the editing reagents to the target cells is one of the crucial and challenging factors for the therapeutic success not only of BEs but also of most Cas-derived genome engineering tools [82]. Selection of the appropriate cargo [DNA/RNA/RNP (ribonucleoprotein)] and the type of delivery vehicle (viral/physical/chemical) remains the most important parameter for achieving successful delivery [82,83]. The size of the BEs also constrains reagent delivery. In addition to DNA-free base-editing strategies (Box 3), the identification of new Cas9 orthologs with smaller sizes would facilitate the delivery of effective therapeutics. A recently characterized smaller Cas protein, CasX, represents one such example [84]. Further development towards modulating the expression of BEs and overcoming the sequence preferences of CBEs would increase their efficacy. For example, base editing of a target genomic locus sometimes has unique requirements for an editing window and PAM compatibility.

CBEs generate more indels, off-target editing, and undesired mutations than do ABEs [25,29,30]. The paramount utility of BEs lies in their ability to install precise base changes, and the generation of undesired base changes (C → A or C → G), bystander and off-target edits, and indels significantly hinders their application in therapeutics. These undesired edits are less problematic for crop improvement because desired edits and indels may be generated on separate alleles that can be fixed through segregation and assortment. A separate off-target edit-evaluation method is needed because the off-target effects of Cas9 and BEs may not be always similar [17]. Although a hypothetical model to design a C → G BE has been proposed [109], HDR will remain the method of choice when there is a need for transversion mutation, changing multiple bases, or installing structural rearrangement. Nevertheless, it is evident that base editing will play a leading role among the many CRISPR-derived technologies for basic research, human therapeutics, and crop improvement.

Outstanding Questions

BEs are not equally efficient at all genomic sites. Despite being accessible to wild-type Cas9, some genomic loci are inaccessible to BEs. How can we engineer BEs to make them suitable for those genomic sites?

Unintended base conversions (e.g., C → A, C → G) within the protospacer region or adjacent region could result in nonsynonymous mutation. Editing of additional C/A bases in the activity window may also give rise to unwanted amino acid alterations in the encoded protein. How can we improve BEs to address these issues while maintaining the efficiency of editing?

In BEs, deaminases are targeted to specific loci by Cas proteins. However, unanticipated interactions of deaminases with ssDNA (created by other cellular enzymes) in the genome cannot be ruled out. What measures should be taken to include such effects in our analysis of off-target mutation?

Currently available BEs can install only transition mutations. There is no known enzyme which can be utilized to develop BEs for transversion mutation. Can directed protein evolution be used to design such types of BEs?

Human DNA repair protein Rev1, a Y family DNA polymerase, is known to insert C opposite to uracil or an abasic site, and plays an important role in C to G transversion mutations during somatic hypermutation. Can Rev1 be utilized to generate a transversion BE?

Will BEs prove to be effective and successful tools for plant *R* gene evolution and for domestication of wild relatives of crop species?

Acknowledgments

Kutubuddin Molla greatly acknowledges the USA-India Educational Foundation (USIEF), New Delhi, and the US Department of State for a Fulbright Nehru Post-Doctoral Fellowship. This work was supported by the National Science Foundation (NSF) Plant Genome Research Program (grant 1740874) and by United States Department of Agriculture (USDA) National Institute of Food and Agriculture (NIFA) Hatch Act Appropriations under project PEN04659 and accession 1016432 to Yinong Yang.

References

1. Komor, A.C. *et al.* (2017) CRISPR-based technologies for the manipulation of eukaryotic genomes. *Cell* 168, 20–36
2. Sternberg, S.H. and Doudna, J.A. (2015) Expanding the biologist's toolkit with CRISPR-Cas9. *Mol. Cell* 58, 568–574
3. Hsu, P.D. *et al.* (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157, 1262–1278
4. Song, F. and Stieger, K. (2017) Optimizing the DNA donor template for homology-directed repair of double-strand breaks. *Mol. Ther. Nucleic Acids* 7, 53–60
5. Bollen, Y. *et al.* (2018) How to create state-of-the-art genetic model systems: strategies for optimal CRISPR-mediated genome editing. *Nucleic Acids Res.* 46, 6435–6454
6. Savic, N. *et al.* (2018) Covalent linkage of the DNA repair template to the CRISPR-Cas9 nuclease enhances homology-directed repair. *eLife* 7, e33761
7. Hess, G.T. *et al.* (2017) Methods and applications of CRISPR-mediated base editing in eukaryotic genomes. *Mol. Cell* 68, 26–43
8. Sun, Y. *et al.* (2016) Engineering herbicide-resistant rice plants through CRISPR/Cas9-mediated homologous recombination of acetolactate synthase. *Mol. Plant* 9, 628–631
9. Komor, A.C. *et al.* (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424
10. Nishida, K. *et al.* (2016) Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* 353
11. Gaudelli, N.M. *et al.* (2017) Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* 551, 464–471
12. Komor, A.C. *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
13. Eid, A. *et al.* (2018) CRISPR base editors: genome editing without double-stranded breaks. *Biochem. J.* 475, 1955–1964
14. Rees, H.A. and Liu, D.R. (2018) Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* 19, 770–788
15. Li, X. *et al.* (2018) Base editing with a Cpf1–cytidine deaminase fusion. *Nat. Biotechnol.* 36, 324–327
16. Rees, H.A. *et al.* (2017) Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. *Nat. Commun.* 8, 15790
17. Kim, D. *et al.* (2017) Genome-wide target specificities of CRISPR RNA-guided programmable deaminases. *Nat. Biotechnol.* 35, 475–480
18. Lin, S. *et al.* (2014) Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife* 3, e04766
19. Cong, L. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823
20. Lindahl, T. (2000) Suppression of spontaneous mutagenesis in human cells by DNA base excision-repair. *Mutat. Res.* 462, 129–135
21. Chen, J. *et al.* (2014) Repair of naturally occurring mismatches can induce mutations in flanking DNA. *eLife*, 3, e02001
22. Alseth, I. *et al.* (2014) Inosine in DNA and RNA. *Curr. Opin. Genet. Dev.* 26, 116–123
23. Liu, Z. *et al.* (2018) Efficient generation of mouse models of human diseases via ABE- and BE-mediated base editing. *Nat. Commun.* 9, 2338
24. Ryu, S.M. *et al.* (2018) Adenine base editing in mouse embryos and an adult mouse model of Duchenne muscular dystrophy. *Nat. Biotechnol.* 36, 536–539
25. Lee, H.K. *et al.* (2018) Targeting fidelity of adenine and cytosine base editors in mouse embryos. *Nat. Commun.* 9, 4804
26. Hua, K. *et al.* (2018) Precise A: T. to G.C. base editing in the rice genome. *Mol. Plant* 11, 627–630
27. Li, C. *et al.* (2018) Expanded base editing in rice and wheat using a Cas9–adenosine deaminase fusion. *Genome Biol.* 19, 59
28. Yan, F. *et al.* (2018) Highly efficient A: T. to G.C. base editing by Cas9n-guided tRNA adenosine deaminase in rice. *Mol. Plant* 11, 631–634
29. Zuo, E. *et al.* (2019) Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science* Published online February 28, 2019. http://dx.doi.org/10.1126/science.aav9973 Published online February 28, 2019
30. Jin, S. *et al.* (2019) Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. *Science* Published online February 28, 2019. http://dx.doi.org/10.1126/science.aaw7166 Published online February 28, 2019
31. Kim, Y.B. *et al.* (2017) Increasing the genome-targeting scope and precision of base editing with engineered Cas9–cytidine deaminase fusions. *Nat. Biotechnol.* 35, 371–376
32. Hua, K. *et al.* (2019) Expanding the base editing scope in rice by using Cas9 variants. *Plant Biotechnol. J.* 17, 499–504
33. Hu, J.H. *et al.* (2018) Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556, 57–63
34. Nishimasu, H. *et al.* (2018) Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* 361, 1259–1262
35. Chatterjee, P. *et al.* (2018) Minimal PAM specificity of a highly similar SpCas9 ortholog. *Sci. Adv.* 4, eaau0766
36. Jakimo, N. *et al.* (2018) A Cas9 with complete PAM recognition for adenine dinucleotides. *bioRxiv* Published online September 27, 2018. http://dx.doi.org/10.1101/429654 Published online September 27, 2018
37. Villiger, L. *et al.* (2018) Treatment of a metabolic liver disease by in vivo genome base editing in adult mice. *Nat. Med.* 24, 1519–1525
38. Yang, L. *et al.* (2016) Engineering and optimising deaminase fusions for genome editing. *Nat. Commun.* 7, 13330
39. Jiang, W. *et al.* (2018) BE-PLUS: a new base editing tool with broadened editing window and enhanced fidelity. *Cell Res.* 28, 855–861
40. Zong, Y. *et al.* (2018) Efficient C-to-T base editing in plants using a fusion of nCas9 and human APOBEC3A. *Nat. Biotechnol.* 36, 950–953
41. Wang, X. *et al.* (2018) Efficient base editing in methylated regions with a human APOBEC3A–Cas9 fusion. *Nat. Biotechnol.* 36, 946–949
42. Gehrke, J.M. *et al.* (2018) An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. *Nat. Biotechnol.* 36, 977–982
43. Tan, J. *et al.* (2019) Engineering of high-precision base editors for site-specific single nucleotide replacement. *Nat. Commun.* 10, 439–439
44. Lu, Y. and Zhu, J.K. (2017) Precise editing of a target base in the rice genome using a modified CRISPR/Cas9 system. *Mol. Plant* 10, 523–525
45. Ren, B. *et al.* (2017) A CRISPR/Cas9 toolkit for efficient targeted base editing to induce genetic variations in rice. *Sci. China Life Sci.* 60, 516–519
46. Zafra, M.P. *et al.* (2018) Optimized base editors enable efficient editing in cells, organoids and mice. *Nat. Biotechnol.* 36, 888–893
47. Tanaka, S. *et al.* (2018) In vivo targeted single-nucleotide editing in zebrafish. *Sci. Rep.* 8, 11423
48. Liang, P. *et al.* (2017) Effective gene editing by high-fidelity base editor 2 in mouse zygotes. *Protein Cell* 8, 601–611
49. Kim, K. *et al.* (2017) Highly efficient RNA-guided base editing in mouse embryos. *Nat. Biotechnol.* 35, 435–437
50. Wang, L. *et al.* (2017) Enhanced base editing by co-expression of free uracil DNA glycosylase inhibitor. *Cell Res.* 27, 1289–1292

51. Koblan, L.W. *et al.* (2018) Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat. Biotechnol.* 36, 843–846
52. Suzuki, K. *et al.* (2016) In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* 540, 144–149
53. Tang, W. *et al.* (2017) Aptazyme-embedded guide RNAs enable ligand-responsive genome editing and transcriptional activation. *Nat. Commun.* 8, 15939
54. Chadwick, A.C. *et al.* (2017) In vivo base editing of PCSK9 (proprotein convertase subtilisin/kexin type 9) as a therapeutic alternative to genome editing. *Arterioscler. Thromb. Vasc. Biol.* 37, 1741–1747
55. Yeh, W.H. *et al.* (2018) In vivo base editing of post-mitotic sensory cells. *Nat. Commun.* 9, 2184
56. Nami, F. *et al.* (2018) Strategies for in vivo genome editing in nondividing cells. *Trends Biotechnol.* 36, 770–786
57. Ma, Y. *et al.* (2016) Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. *Nat. Methods* 13, 1029–1035
58. Hess, G.T. *et al.* (2016) Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. *Nat. Methods* 13, 1036–1042
59. Kuscu, C. *et al.* (2017) CRISPR-STOP: gene silencing through base-editing-induced nonsense mutations. *Nat. Methods* 14, 710–712
60. Billon, P. *et al.* (2017) CRISPR-mediated base editing enables efficient disruption of eukaryotic genes through induction of STOP codons. *Mol. Cell* 67, 1068–1079
61. Gapinske, M. *et al.* (2018) CRISPR-SKIP: programmable gene splicing with single base editors. *Genome Biol.* 19, 107
62. Zhang, S. *et al.* (2018) Base-specific mutational intolerance near splice sites clarifies the role of nonessential splice nucleotides. *Genome Res.* 28, 968–974
63. Yuan, J. *et al.* (2018) Genetic modulation of RNA splicing with a CRISPR-guided cytidine deaminase. *Mol. Cell* 72, 380–394
64. Liu, Z. *et al.* (2018) Highly efficient RNA-guided base editing in rabbit. *Nat. Commun.* 9, 2717
65. Laloum, T. *et al.* (2018) Alternative splicing control of abiotic stress responses. *Trends Plant Sci.* 23, 140–150
66. Xue, C. *et al.* (2018) Manipulating mRNA splicing by base editing in plants. *Sci. China Life Sci.* 61, 1293–1300
67. Jinek, M. *et al.* (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821
68. Qi, L.S. *et al.* (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152, 1173–1183
69. Panda, D. *et al.* (2018) DNA as a digital information storage device: hope or hype? *3 Biotech* 8, 239
70. Tang, W. and Liu, D.R. (2018) Rewritable multi-event analog recording in bacterial and mammalian cells. *Science* 360, eaap8992
71. Farzadfard, F. *et al.* (2018) Single-nucleotide-resolution computing and memory in living cells. *bioRxiv* Published online February 16, 2018. http://dx.doi.org/10.1101/263657 Published online February 16, 2018
72. Farzadfard, F. and Lu, T.K. (2018) Emerging applications for DNA writers and molecular recorders. *Science*, 361, 870–875
73. Zeng, Y. *et al.* (2018) Correction of the Marfan syndrome pathogenic FBN1 mutation by base editing in human cells and heterozygous embryos. *Mol. Ther.* 26, 2631–2637
74. Rossidis, A.C. *et al.* (2018) In utero CRISPR-mediated therapeutic editing of metabolic genes. *Nat. Med.* 24, 1513–1518
75. Zhao, H. *et al.* (2005) Neighboring-nucleotide effects on the mutation patterns of the rice genome. *Genomics Proteomics Bioinformatics* 3, 158–168
76. Shimatani, Z. *et al.* (2017) Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. *Nat. Biotechnol.* 35, 441–443
77. van Schie, C.C. and Takken, F.L. (2014) Susceptibility genes 101: how to be a good host. *Annu. Rev. Phytopathol.* 52, 551–581
78. Moscou, M.J. and Bogdanove, A.J. (2009) A simple cipher governs DNA recognition by TAL effectors. *Science* 326, 1501
79. Jacob, P. *et al.* (2018) Translational research: exploring and creating genetic diversity. *Trends Plant Sci.* 23, 42–52
80. Farnham, G. and Baulcombe, D.C. (2006) Artificial evolution extends the spectrum of viruses that are targeted by a disease-resistance gene from potato. *Proc. Natl. Acad. Sci. U. S. A.* 103, 18828–18833
81. Harris, C.J. *et al.* (2013) Stepwise artificial evolution of a plant disease resistance gene. *Proc. Natl. Acad. Sci. U. S. A.* 110, 21189–21194
82. Liu, C. *et al.* (2017) Delivery strategies of the CRISPR-Cas9 gene-editing system for therapeutic applications. *J. Control. Release* 266, 17–26
83. Glass, Z. *et al.* (2018) Engineering the delivery system for CRISPR-based genome editing. *Trends Biotechnol.* 36, 173–185
84. Liu, J.J. *et al.* (2019) CasX enzymes comprise a distinct family of RNA-guided genome editors. *Nature* 566, 218–223
85. Woo, J.W. *et al.* (2015) DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat. Biotechnol.* 33, 1162–1164
86. Cox, D.B.T. *et al.* (2017) RNA editing with CRISPR-Cas13. *Science* 358, 1019–1027
87. Park, D.S. *et al.* (2017) Targeted base editing via RNA-guided cytidine deaminases in *Xenopus laevis* embryos. *Mol. Cell* 40, 823–827
88. Zhang, Y. *et al.* (2017) Programmable base editing of zebrafish genome using a modified CRISPR-Cas9 system. *Nat. Commun.* 8, 118
89. Zheng, K. *et al.* (2018) Highly efficient base editing in bacteria using a Cas9–cytidine deaminase fusion. *Commun. Biol.* 1, 32
90. Li, Y. *et al.* (2018) Programmable single and multiplex base-editing in *Bombyx mori* using RNA-guided cytidine deaminases. *G3 (Bethesda)* 8, 1701–1709
91. Ma, Y. *et al.* (2018) Highly efficient and precise base editing by engineered dCas9-guide tRNA adenosine deaminase in rats. *Cell Discov.* 4, 39
92. Lee, J.K. *et al.* (2018) Directed evolution of CRISPR-Cas9 to increase its specificity. *Nat. Commun.* 9, 3048
93. Yang, L. *et al.* (2018) Increasing targeting scope of adenosine base editors in mouse and rat embryos through fusion of TadA deaminase with Cas9 variants. *Protein Cell* 9, 814–819
94. Qin, W. *et al.* (2018) Precise A*T to G*C base editing in the zebrafish genome. *BMC Biol.* 16, 139
95. Chen, Y. *et al.* (2017) CRISPR/Cas9-mediated base-editing system efficiently generates gain-of-function mutations in *Arabidopsis*. *Sci. China Life Sci.* 60, 520–523
96. Li, J. *et al.* (2017) Generation of targeted point mutations in rice by a modified CRISPR/Cas9 system. *Mol. Plant* 10, 526–529
97. Zong, Y. *et al.* (2017) Precise base editing in rice, wheat and maize with a Cas9–cytidine deaminase fusion. *Nat. Biotechnol.* 35, 438–440
98. Kang, B.C. *et al.* (2018) Precision genome engineering through adenine base editing in plants. *Nat. Plants* 4, 427–431
99. Hua, K. *et al.* (2018) Expanding the base editing scope in rice by using Cas9 variants. *Plant Biotechnol. J.* 17, 499–504
100. Endo, M. *et al.* (2018) Genome editing in plants by engineered CRISPR-Cas9 recognizing NG PAM. *Nat. Plants* 5, 14–17
101. Barnes, D.E. and Lindahl, T. (2004) Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu. Rev. Genet.* 38, 445–476
102. Zong, Y. *et al.* (2018) Efficient C-to-T base editing in plants using a fusion of nCas9 and human APOBEC3A. *Nat. Biotechnol.* 36, 950–953

103. Grunebaum, E. *et al.* (2013) Recent advances in understanding and managing adenosine deaminase and purine nucleoside phosphorylase deficiencies. *Curr. Opin. Allergy Clin. Immunol.* 13, 630–638

104. Abudayyeh, O.O. *et al.* (2016) C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* 353, aaf5573

105. Zuris, J.A. *et al.* (2015) Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat. Biotechnol.* 33, 73–80

106. Liang, X. *et al.* (2015) Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *J. Biotechnol.* 208, 44–53

107. Sasaguri, H. *et al.* (2019) Introduction of pathogenic mutations into the mouse Psen1 gene by base editor and target-AID. *Nat. Commun.* 9, 2892

108. Wu, Z. *et al.* (2010) Effect of genome size on AAV vector packaging. *Mol. Ther.* 18, 80–86

109. Gajula, K.S. (2019) Designing an elusive C•G→G•C CRISPR base editor. *Trends Biochem. Sci.* 44, 91–94