Trends in Genetics

CellPress OPEN ACCESS

Feature Review

DNA base editing in nuclear and organellar genomes

Junjie Tan [1,3], Joachim Forner [2,3], Daniel Karcher,[2] and Ralph Bock [2,*]

Genome editing continues to revolutionize biological research. Due to its simplicity and flexibility, CRISPR/Cas-based editing has become the preferred technology in most systems. Cas nucleases tolerate fusion to large protein domains, thus allowing combination of their DNA recognition properties with new enzymatic activities. Fusion to nucleoside deaminase or reverse transcriptase domains has produced base editors and prime editors that, instead of generating double-strand breaks in the target sequence, induce site-specific alterations of single (or a few adjacent) nucleotides. The availability of protein-only genome editing reagents based on transcription activator-like effectors has enabled the extension of base editing to the genomes of chloroplasts and mitochondria. In this review, we summarize currently available base editing methods for nuclear and organellar genomes. We highlight recent advances with improving precision, specificity, and efficiency and discuss current limitations and future challenges. We also provide a brief overview of applications in agricultural biotechnology and gene therapy.

Highlights

With the development of base editors, genome editing tools have become available that introduce specific point mutations in a targeted manner.

Cytosine base editors and adenine base editors catalyze site-specific nucleoside deamination to introduce C-to-T and A-to-G mutations, respectively.

Prime editing uses fusions of Cas proteins to reverse transcriptases for the guide RNA-templated copying of point mutations, or small insertions or deletions into the target DNA.

TALE- and TALEN-based editing tools have been developed to introduce point mutations into the genomes of the two DNA-containing cell organelles, plastids (chloroplasts) and mitochondria.

Introduction

Genome editing technologies take advantage of sequence-programmable nucleases to specifically modify a target region in the genome. Three major genome editing technologies have been developed: **zinc finger nucleases (ZFNs)** (see Glossary), **transcription activator-like effector nucleases (TALENs)**, and **clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas)** [1–3]. While ZFNs and TALENs recognize the target sequence by DNA-binding protein modules, CRISPR/Cas systems rely on DNA recognition by small RNAs, so-called single guide RNAs (sgRNAs), via complementary base pairing. Although structurally somewhat more complex due to the requirement for both a protein and an RNA component, CRISPR/Cas systems have proven to be superior in design simplicity and efficiency and quickly became the most widely used genome editing technique for site-specific DNA manipulations in nearly all organisms that are genetically transformable [4–6].

Classical genome editing is achieved by tailor-made editing reagents introducing double-strand break(s) (DSB) in the target site that subsequently are repaired by one of the cellular DNA repair mechanisms [7] (Figure 1). In the cells of most eukaryotes, three main pathways for DSB repair exist: **non-homologous end joining (NHEJ)**, **microhomology-mediated end joining (MMEJ)**, and **homology-directed repair (HDR)** [8,9]. NHEJ and MMEJ are usually more efficient than HDR, but tend to be error-prone, making the outcome of the repair process largely unpredictable, although, recently, progress has been made with predicting the outcome of at least some types of insertions and/or deletions (InDels) [10]. DSB repair by these two pathways can produce insertions, deletions, translocations, or other DNA rearrangements at the cleavage sites [8]. By contrast, HDR exhibits high repair precision, but in most organisms displays very low efficiency [10], presumably because of its dependence on the presence of a donor template

[1]State Key Laboratory of Crop Genetics and Germplasm Enhancement, Innovation Center for Genome Editing and Engineering, Nanjing Agricultural University, Nanjing, 210095, China
[2]Max-Planck-Institut für Molekulare Pflanzenphysiologie, Am Mühlenberg 1, D-14476 Potsdam-Golm, Germany
[3]These authors contributed equally to this article

*Correspondence:
rbock@mpimp-golm.mpg.de (R. Bock).

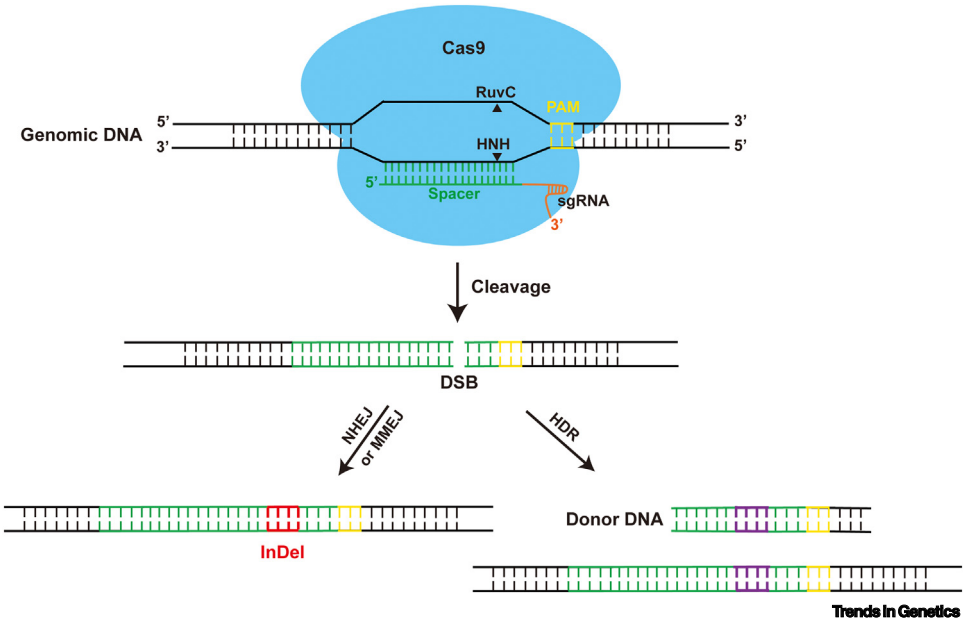

Figure 1. Clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9)-mediated genome editing. The Cas9-single guide RNA (sgRNA) ribonucleoprotein complex recognizes the target site in the genomic DNA through complementarity with the spacer sequence (green) of the sgRNA. Subsequently, the HNH and RuvC nuclease domains of the Cas9 protein cleave the two DNA strands, creating a double-strand break (DSB) approximately three nucleotides upstream of the **protospacer adjacent motif (PAM)** sequence (yellow). DSB repair can occur by one of three pathways: non-homologous end joining (NHEJ), microhomology-mediated end joining (MMEJ), and homology-directed repair (HDR). NHEJ and MMEJ repair are error prone and can result in more or less random insertions or deletions (InDels) at the cleavage site. By contrast, HDR repairs the DSB with high precision by utilizing a donor DNA as template. The donor DNA contains homology arms that flank the DSB site and the mutations to be introduced (purple).

(Figure 1) and the low abundance and/or activity of the recombination machinery in most cell types and developmental stages [9].

Programmable nucleases such as Cas9 induce DSBs in the target region, which are predominantly repaired by NHEJ. Consequently, the end products of the repair process often harbor larger DNA alterations that cause disruption of the target gene [3]. Thus, although CRISPR/Cas9 represents a powerful tool for the analysis of gene functions by reverse genetics, it is not suitable for precise sequence manipulations of genomic DNA, as required, for example, in gene therapy and crop breeding. This serious deficit has triggered efforts to develop a new type of genome editing reagents, dubbed base editors (BEs), that facilitate the precise replacement of individual bases in the DNA without generating DSBs at the target site and without the need to provide exogenous repair templates [11–15]. Catalytically inactive (deactivated) Cas9 proteins (dCas9) or Cas9 nickases, that introduce single-strand breaks rather than cutting through the DNA double strand, largely prevent the introduction of InDels (Figure 1), while fully retaining the DNA recognition properties conferred by the sgRNA [11,15]. This feature has enabled construction of fusion proteins that endow CRISPR/Cas systems with new enzymatic properties, including nucleotide-modifying activities.

The vast majority of the genetic material of eukaryotic cells resides in the nuclear genome. However, essential information is also encoded in the much smaller genomes of the DNA-containing organelles: the **mitochondria** and, in plant cells, the **plastids** (chloroplasts). The manipulation of

Glossary

Adenosine deaminase: an enzyme that catalyzes the deamination of adenosine, thus resulting in conversion to inosine.

***Agrobacterium*-mediated transformation:** a method to integrate foreign DNA into the nuclear genome of plant cells, which exploits the natural ability of the soil bacterium *Agrobacterium tumefaciens* to introduce a defined segment of its DNA (the T-DNA) into the nucleus of the host cell. The T-DNA can be replaced by a DNA sequence of interest that then is faithfully delivered to the plant nucleus by the bacterium.

Base excision repair (BER): a cellular mechanism that repairs short DNA sequences in which one or more bases have been chemically altered by oxidation, deamination, or alkylation. A short sequence stretch containing the modified base(s) is excised followed by repair synthesis.

Circularly permuted: reorganization of the topology of the primary sequence of a protein. The original N and C termini of the protein are covalently linked and the protein is divided at different positions.

Clustered regularly interspaced short palindromic repeats (CRISPR): an array of DNA sequences derived from infectious agents (viruses or plasmids) and arranged as short direct repeats in the genomes of many bacteria and archaea.

CRISPR-associated (Cas): a CRISPR-related gene often located adjacent to the CRISPR locus. Its gene product typically uses transcribed CRISPR sequence information to recognize and cleave invading DNA elements.

CRISPR RNA (crRNA): noncoding RNA generated by transcription of the CRISPR locus. It guides the Cas nuclease complex to its specific cleavage site.

Cytidine deaminase: an enzyme catalyzing the deamination of cytidine, thus converting it to uridine.

Gene drive: a naturally occurring or technically engineered situation in which a given piece of DNA (e.g., a specific allele of a gene) has a direct selective advantage over other alleles already at the DNA level (i.e., independent of phenotypic selection), thus promoting transmission to the progeny. Gene drives can occur at the level of the individual cell (e.g., in the highly polyploid

the genomes of the cytoplasmic organelles continues to be challenging and direct methods for organellar genome transformation are available only in a limited number of species. Proteins expressed from the nuclear genome can be imported into the organelles, but, with the exception of some tRNAs that can be imported into mitochondria, the import of RNA molecules is not generally possible. Consequently, in the absence of delivery methods for sgRNAs, CRISPR/Cas systems currently cannot be directed towards organellar genomes. However, recently, protein-only genome editing reagents such as TALENs have been successfully engineered to facilitate base editing in mitochondrial and plastid genomes.

In this review, we summarize currently available base editing methods and their various embodiments to enhance the efficiency and increase the accuracy of site-specific sequence modification. In addition to classical base editing methods that rely on nucleoside deaminase fusions, we also cover related technologies that use genome editing reagents to achieve site-directed sequence changes in nuclear and organellar genomes. We discuss current limitations and future directions for the development of improved base editing tools and highlight selected applications of DNA base editing in gene therapy and agricultural biotechnology.

The CRISPR/Cas9 system: a brief overview

The CRISPR/Cas9 system is a naturally occurring defense mechanism of bacteria. Its biological function is to provide protection against invading phages, plasmids, and possibly other types of potentially harmful environmental DNA [3,16]. The system represents a ribonucleoprotein, in which the RNA component confers the sequence specificity by complementary base pairing with the target DNA (Figure 1). In natural CRISPR/Cas9 systems, the RNA component is comprised of two small RNA molecules: the **CRISPR RNA (crRNA)** transcribed from the CRISPR locus, and the ***trans*-activating crRNA (tracrRNA)** that is involved in maturation of the crRNA [17]. The crRNA:tracrRNA duplex, sometimes also referred to as dual guide RNA, can be engineered into an sgRNA [3] (Figure 1), thus greatly simplifying the design of synthetic systems with new target specificities. The protein component is an endonuclease (Cas9) composed of multiple domains, of which HNH- and RuvC-like domains provide the nuclease activities that cleave the two strands of the target DNA (Figure 1) [3,16,18]. Cas9 cuts the target DNA in a series of coordinated steps. First, Cas9, in complex with the guide RNA (sgRNA), recognizes a so-called **protospacer adjacent motif (PAM)** sequence (e.g., 5′-NGG-3′) via its PAM-interacting domain. The sgRNA then base pairs with the first 10–12 nucleotides (referred to as seed sequence) of the target DNA [19–21] (Figure 1). Annealing of the seed sequence with the sgRNA makes the target DNA more accessible, thus facilitating the extended annealing of the sgRNA with the target DNA strand. This process results in the formation of an **R-loop**, followed by cleavage of both DNA strands by the Cas9 nuclease domains (Figure 1).

The versatility of CRISPR/Cas9 systems was further increased by engineering Cas9 variants with altered PAM specificities, thus expanding the genome-targeting scope and making more target sequences accessible to genome editing [22]. For example, Cas9-VQR recognizes the PAM sequence 5′-NGA-3′, and Cas9-VRER recognizes 5′-NGCG-3′ [23]. Other examples are the xCas9 variant that is compatible with NG, GAA, and GAT PAMs and was generated through directed evolution, and the rationally designed Cas9-NG that displays relaxed PAM requirement by recognizing 5′-NG-3′ [24,25]. Recently, variants SpG and SpRY were engineered for nearly PAM-independent target recognition [26]. Although relaxed PAM requirement can come at the cost of reduced editing efficiency, the availability of a large set of Cas9 variants with altered PAM specificities has significantly expanded the possibilities of genome editing by largely removing sequence context-related restrictions on target site selection [27,28].

plastids and mitochondria) or at the population level. Underlying mechanisms can include allele conversion, preferential replication, decreased degradation, increased probability of entry into the egg cell during female gametogenesis, or inactivation of sperm cells not carrying the drive allele after male gametogenesis.

Homochondriomy: the state in which all copies of the mitochondrial genome in a given cell are identical (in contrast to heterochondriomy, where there are at least two different genome types present).

Homologous recombination (HR): a pathway to repair DNA double-strand breaks, which requires the presence of at least one intact homologous copy of the broken DNA. The broken DNA strand(s) invade the intact homologous DNA and use it as template for repair synthesis (see also 'homology-directed repair').

Homology-directed repair (HDR): a highly accurate DNA double-strand break repair pathway that uses a homologous DNA donor sequence or the sister chromatid for repairing DSBs. Often used synonymously with homologous recombination repair (HRR).

Homoplasmy: the state in which all copies of the plastid genome in a given cell are identical (in contrast to heteroplasmy, where there are at least two different genome types present).

Lipofection: liposome-based transfection, a method to deliver DNA into cells where the cargo DNA is encapsulated by a lipid bilayer.

Microhomology-mediated end joining (MMEJ): a DNA repair pathway that uses microhomologous sequences to repair double-strand breaks. At the repaired site, deletions, translocations, or inversions can occur.

Mitochondria: DNA-bearing organelles in the cytoplasm of eukaryotic cells, which are originated from the endosymbiotic uptake of an α-proteobacterium. One of their main functions is to provide energy via oxidative phosphorylation.

Non-homologous end joining (NHEJ): a DNA repair pathway that directly reseals DNA breaks by re-ligation, without the use of a repair template. NHEJ is error-prone and can result in insertions or deletions at the repair site.

Plastids: DNA-bearing organelles in plant cells, which are derived from a cyanobacterial ancestor acquired by endosymbiosis more than a billion years ago. Plastids occur in several

CRISPR/Cas-derived BEs

BEs change the identity of individual nucleobases at the target site in the genome without creating DSBs and/or requiring DNA repair templates [11–15,29]. Current BEs generally consist of a catalytically inactive Cas nuclease protein (dCas), or a Cas nickase (nCas), fused to a nucleoside deaminase enzyme [11,15,29]. Based on the two types of nucleoside deaminases that can be used, two classes of BEs are distinguished: cytosine base editors (CBEs) and adenine base editors (ABEs) [11,15]. CBEs carry a **cytidine deaminase** fused to Cas9, which causes C-to-T transitions by deaminating deoxycytidine to deoxyuridine, a nucleoside that has similar base pairing properties as deoxythymidine. In ABEs, an **adenosine deaminase** is tethered to the Cas9 protein (Figure 2), which catalyzes the oxidative deamination of deoxyadenosine to deoxyinosine. Because deoxyinosine mimics deoxyguanosine (in that it preferentially base pairs with deoxycytidine), this conversion causes A-to-G transitions. Uracil and hypoxanthine (the nucleobase of the nucleoside inosine) are both alien bases in DNA, which normally are recognized and removed by the **base excision repair (BER)** pathway. This makes the fixation of the mutation introduced by BEs dependent on improper repair or lack of repair prior to DNA replication (Figure 2) and poses a serious limitation on the efficiency of base editing in DNA (see later).

CBEs and ABEs can induce all four base transitions (C to T, G to A, A to G, and T to C) when considering both DNA strands. Unlike conventional CRISPR/Cas-mediated genome editing, in which the editing outcome is largely unpredictable and depends on the cellular DNA repair mechanism utilized to fix the DSB (Figure 1), deaminase-type BEs represent precise instruments for the targeted alteration of the identity of single bases in DNA (Figure 2). Importantly, BEs do not require provision of a nucleic acid template or any additional exogenous or endogenous protein factors for the nucleotide modification to occur [13,14]. However, the fixation probability of the mutation depends on the activity of DNA repair, replication, and recombination pathways (Figure 2).

BEs have been successfully applied in a wide range of organisms, including human cell cultures, experimental animals, various microbes, and plants [14,30–46]. The possibility to precisely manipulate single bases in genomes offers great potential for curing hereditary diseases. Approximately 30% of the currently annotated clinically relevant human disease variants represent single base transitions that, potentially, could be targeted by appropriately designed BEs [47]. Likewise, many **quantitative trait loci (QTLs)** that underlie phenotypic variation of complex agronomic traits such as crop yield and abiotic stress tolerance are determined by single nucleotide polymorphisms that also could be introduced into elite varieties by base editing. Thus, BEs also promise to revolutionize plant breeding [46,48,49].

Strategies to improve the efficiency of deaminase-based BEs

In the first nucleoside deaminase-type BE, a member of the APOBEC1 family of cytidine deaminases was fused to Cas9 to create the first-generation BE, BE1 [11]. APOBEC1 enzymes catalyze the hydrolysis of cytosine to uracil in polydeoxyribonucleotides. Uracil is read as thymine by replication enzymes (DNA polymerases), but uracil residues in DNA are efficiently detected as sites of oxidative damage and rapidly excised by the enzyme uracil DNA N-glycosylase (UNG). While useful to prevent accumulation of mutations from spontaneously occurring oxidative deamination of cytosine, the efficient removal of uracil from genomic DNA greatly reduces the fixation efficiency of C-to-T transitions caused by CBEs [11,50]. The addition of a uracil glycosylase inhibitor (UGI) to the BE1 fusion protein (generating the second-generation BE, BE2) substantially increased the efficiency of C-to-T base editing by inhibiting UNG-mediated BER [11,51,52]. UGI is a small protein (of 9.5 kDa) from *Bacillus subtilis* bacteriophage PBS1 and its fusion to the C terminus of CBEs is well tolerated.

differentiation forms, including chloroplasts in which photosynthesis takes place.

Protospacer: a DNA sequence in an invading DNA element, which has become integrated into a bacterial CRISPR locus and can be transcribed into a crRNA.

Protospacer adjacent motif (PAM): a short DNA sequence comprised of two to five nucleotides, which is often found adjacent to the protospacer. The PAM is specifically recognized by the corresponding Cas protein as part of the sequence-specific target site recognition for endonucleolytic cleavage.

Quantitative trait locus (QTL): a genomic locus that contributes to a highly polygenic trait (quantitative trait).

R-loop: a nucleic acid structure consisting of three strands: a DNA/RNA hybrid and the DNA strand displaced by the RNA.

***Trans*-activating crRNA (tracrRNA):** an additional noncoding RNA found in type II CRISPR/Cas9 systems, which is partially complementary to the pre-crRNA. After processing by RNase III, the RNA–RNA duplex guides the Cas9 nuclease complex to its target site.

Transcription activator-like effector (TALE): a class of proteins secreted by certain plant-pathogenic bacteria (e.g., *Xanthomonas citri*, the causative agent of citrus canker). TALEs bind to specific DNA target sequences in the plant nucleus and usually activate transcription of these loci (to the pathogen's benefit). Repeats of 34 amino acids each facilitate target-specific DNA binding; an activation domain promotes transcription.

Transcription activator-like effector nuclease (TALEN): engineered site-specific DNA endonucleases based on TALEs (see earlier), which are created by replacing the TALE activation domain by the FokI restriction endonuclease domain.

Zinc finger nuclease (ZFN): a protein motif (named for its propensity to coordinate zinc ions) that is frequently found as a DNA-binding domain in transcription factors. One zinc finger unit recognizes a target sequence of three base pairs in the DNA. When coupled to the FokI endonuclease domain, ZFNs can be created analogously to TALENs.

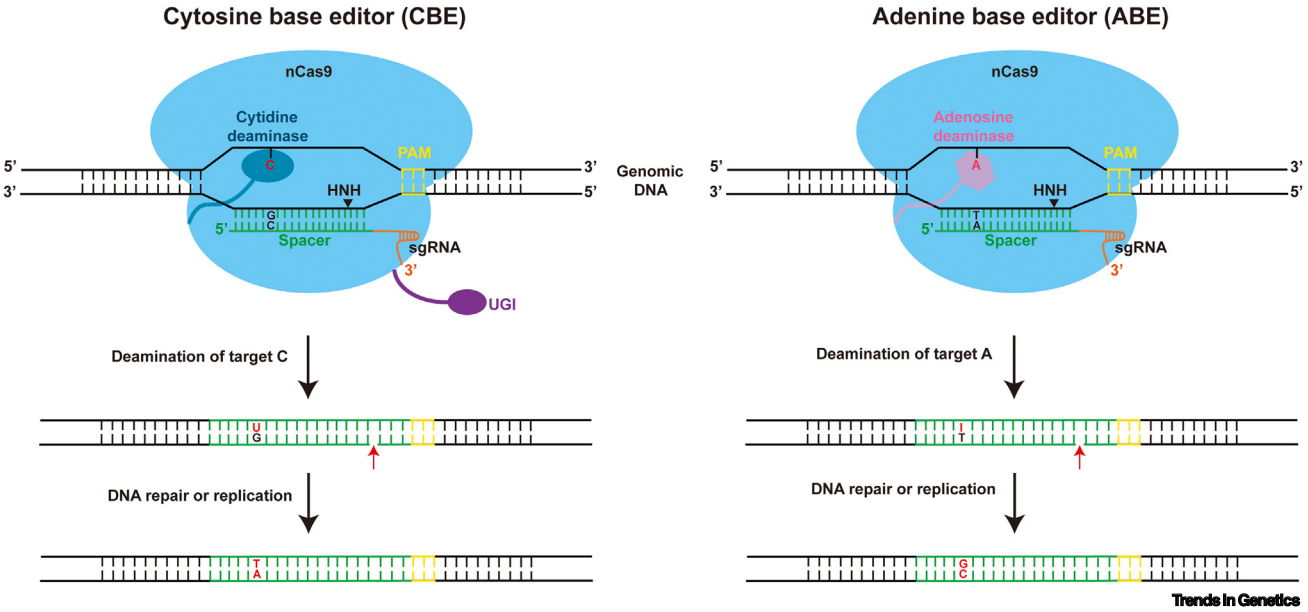

Figure 2. Clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas)-based DNA base editing. Base editing with cytosine base editors (CBEs, left) and adenine base editors (ABEs, right) is illustrated schematically. A CBE consists of a Cas9 nickase (nCas9), a cytidine deaminase (dark blue), and a uracil DNA glycosylase inhibitor (UGI, purple). The UGI protects the Us arising from C deamination by preventing their excision by uracil DNA glycosylase (UDG). The nick in the unedited strand is recognized by the cellular mismatch repair (MMR) pathway that then can use the edited strand as a repair template to form a U:A base pair, which can become fixed as a T:A base pair by DNA repair and/or DNA replication. ABEs consist of nCas9 and an adenosine deaminase (pink) that mediates deamination of adenosine (A) to inosine (I). Inosine is recognized as guanosine (G) upon DNA replication, thus resulting in fixation of an A-to-G mutation.

A further improvement was achieved by replacing dCas9 with the nickase nCas9, as done in the third-generation CBE BE3 and subsequently developed ABEs. The nickase introduces a single-strand break into the non-deaminated DNA strand. The break directs the cellular mismatch repair (MMR) system to the unedited strand and, in this way, stimulates the use of the deaminated strand as repair template [11,15]. Fusion of a second UGI domain to the C terminus of BE3 further increased base conversion efficiency and enhanced product purity [53,54] (fourth-generation CBEs; BE4). Further improvements in BE activity came from: (i) modification of nuclear localization signals and adjustments in codon usage, yielding the BE4max and ABEmax variants [35,55]; (ii) incorporation of a single-stranded DNA (ssDNA)-binding protein domain into CBEs [56]; and (iii) application of directed evolution approaches to generate the eighth-generation ABEs, ABE8s and ABE8e, that display substantially improved activity (with ABE8e catalyzing DNA deamination up to ~1100-fold faster than the previous version ABE7.10) [57–59]. Together, these improvements provided substantially increased editing efficiencies in both mammalian cells and plants [35,48,60].

Engineering BEs for altered activity windows

In the course of complex formation between BE fusion protein, sgRNA, and target DNA, an R-loop is formed by annealing of the sgRNA with the complementary DNA strand (Figure 2). The displaced ssDNA provides the substrate upon which the deaminase of the BE can act. The sequence stretch of the displaced strand that is accessible to the deaminase domain of the fusion protein is referred to as ‘activity window’ or ‘base editing window’ [11,20,61]. Binding of the fusion protein to the target DNA sequence masks PAM-adjacent nucleotides, so that the activity window of current BEs is generally located distal to the PAM [11,25,29,61]. For example, the activity window of the canonical BE3 is approximately five nucleotides, covering the positions 4 to 8 (counting from the PAM-distal end of the 20-bp spacer, with the PAM corresponding to positions 21 to 23), equivalent to nucleotide positions –13 to –17 relative to the PAM sequence [11,15].

The activity window of BEs can be influenced by a number of factors, including the choice of the deaminase and the Cas9 protein and the conformation of the DNA [11,14,15,62–64]. As base editing within the R-loop depends on molecular interaction between the tethered deaminase and the substrate nucleotides, size and three-dimensional structure of the deaminase and Cas9 domains of the fusion protein affect location and width of the editing window [62]. Systematic testing of different deaminase domains has produced BEs with different activity windows. For example, the CBE BE3 containing the deaminase APOBEC1 from *Rattus norvegicus* has an activity window from C_4 to C_8 [11], meaning that any deoxycytidine present between position 4 and 8 of the spacer (i.e., positions –13 to –17 upstream of the PAM) can become deaminated, albeit at different efficiency. By contrast, a BE constructed with another cytidine deaminase, pmCDA1 from sea lamprey, called target activation-induced cytidine deaminase, or Target-AID, has an activity window from C_{-16} to C_{-19} relative to the PAM (corresponding to positions 2 to 5 relative to the PAM-distal end of the spacer) [29]. Various cytidine deaminases have been tested to diversify CBEs, including mammalian APOBEC family proteins such as AID [65–68], APOBEC1 and APOBEC3 (A3A-H) proteins [69–74], and lamprey-derived CDA1 family proteins such as CDA1 [29,63,64,75–77]. Simultaneous incorporation of multiple (identical or different) deaminase domains into a BE fusion protein broadens the accessibility of substrate nucleotides in the R-loop, thus widening the base editing window [78], diversifying editing reactions (i.e., by combining a cytidine deaminase domain with an adenosine deaminase domain) and/or increasing editing efficiency [79–83]. Cytidine deaminases also exhibit certain preferences related to the sequence composition of their substrate nucleic acid, which can further influence the width of the activity window of CBEs, but also their activity on different targets. For example, APOBEC1 deaminases display nearest-neighbor preferences in the order TC ≥ CC ≥ AC > GC [11]. By contrast, the APOBEC3G (A3G) deaminase exhibits the highest activity on the second C in CC dinucleotide motifs [73,74]. BEs constructed with CDA, AID, and A3A deaminases appear to be less dependent on the sequence context of the substrate DNA, likely due to the generally higher activity of these deaminases [29,65,66,69–71].

Naturally occurring or artificially engineered Cas protein variants, in addition to varying in their PAM sequence requirements, can also offer different editing windows. For example, *Staphylococcus aureus* Cas9 (SaCas9)-based BEs edit a broader editing window, typically between positions 3 and 12 for CBEs and positions 4 to 12 for ABEs (relative to the PAM-distal end of the spacer) [84]. Cas12a-based ABEs and CBEs edit in an activity window between positions 8 and 13 [85], and a minimal Cas12f variant (CasMINI)-based ABE has a very narrow window of only two nucleotides (positions 3 and 4 relative to the PAM) [86]. **Circularly permuted** Cas9 variants (Cas9-CP) alter the localization of the deaminase domain in the R-loop and, in this way, can enable more efficient base conversion in a broader editing window [87,88].

The activity window can also be modified by the orientation of the domains within the BE fusion protein (Figure 2), the properties of the linker connecting them, and the size of the deaminase domain. For example, BEs with CDA1 fused to either the N terminus or the C terminus of nCas9 (nCDA1-BE3 and cCDA1-BE3) have different activity windows [63]. In general, deaminases fused to the N terminus of nCas9 exhibit broader editing windows [11,63,75], likely because this orientation provides easier access to the ssDNA within the R-loop. Interestingly, recent studies report that deaminases can be embedded into the Cas9 protein (rather than fusing them to the N or C terminus of Cas9) and the resulting BEs have different editing windows and activities, depending on the deaminase position within the Cas9 protein [89,90].

The linkers that connect the different domains of BE fusion proteins (Figure 2) can also influence location and width of the editing window as well as editing activity. For example, reducing the

linker length in BE3 or replacing its flexible linker by more rigid linkers (containing the secondary amino acid proline, whose side chain is doubly connected to the protein backbone, thus restricting the flexibility of the polypeptide chain) results in reduced size of the editing window [63,64]. Presumably, closer spacing between the deaminase domain and the Cas9 domain of the fusion protein results in a more precise positioning of the deaminase on the target sequence. By replacing canonical linkers with alternative linkers (varying in length and stiffness) and removing nonessential sequences from the CDA1 deaminase, a series of BEs with narrower activity windows could be obtained [63,64].

Since base editing depends on direct interaction between the catalytic domain of the deaminase and the substrate nucleotide, engineering deaminases with modified catalytic activity can also result in altered editing windows. For example, variants of APOBEC1 (YE1, YE2, and YEE) were created by mutating residues involved in the catalytic site or the substrate-binding domain of APOBEC1, resulting in narrower editing windows [84]. Similarly, an introduced point mutation in APOBEC3A (N57G) generated a CBE (eA3A) with preference for TCR motifs [69], whereas an amino acid exchange in ABE7.10 (F148A) resulted in a BE with a narrower editing window [91].

Engineered BEs for improved specificity and precision

The editing activity of all currently available BEs is not strictly limited to the intended target site, but can also occur at unwanted sites elsewhere in the genome. This so-called off-target DNA editing can occur in either an sgRNA-dependent or sgRNA-independent manner. sgRNA-dependent off-target editing happens when the editing complex is guided by the sgRNA to sequences with similarity to the spacer (Figure 1) [92–94]. This problem can be largely avoided by using high-fidelity Cas variants [24,43,95–98], careful sgRNA design, use of truncated sgRNAs [99,100], and/or delivering the BE as ribonucleoprotein particle (RNP) preassembled *in vitro* [101,102].

Random off-target mutations in the genome are induced in an sgRNA-independent way, especially upon high-level expression of the editing complex. This can lead to deamination of accessible nucleotides preferentially in coding regions, where melting of the double-stranded DNA by the transcribing RNA polymerase complex generates accessible single-stranded regions. This phenomenon is predominantly observed in CBEs and not so much in ABEs, probably because the adenosine deaminase used in ABEs is derived from an RNA editing enzyme that does not possess high affinity to DNA [103,104]. sgRNA-independent off-target editing in the genome can be reduced by engineering appropriate deaminase variants. For example, the engineered APOBEC1-based BE variant YE1-BE3 has substantially reduced off-target activity [105,106]. By screening 153 cytidine deaminases and further optimization for superior on-target and reduced off-target DNA editing, four cytidine deaminase variants (AmAPOBEC1, PpAPOBEC1, RrA3F, and SsAPOBEC3B) with minimized sgRNA-independent off-target editing activity were identified [107]. BEs with properly internally embedded deaminases into the Cas protein also show reduced genome-wide off-target editing activity [89,90]. In general, BEs with narrow editing windows exhibit lower genome-wide off-target editing, because a narrow editing window means that fewer nucleotides are accessible to the BE at both on-target and off-target sites [64,105].

Both CBEs and ABEs can also induce RNA editing in an sgRNA-independent manner, because of their affinity to RNA, which in the case of ABEs, is the natural substrate of the deaminase [91,108–110]. Although an RNA off-target mutation has a limited lifetime and ultimately disappears when the edited RNA molecule is turned over, the frequent induction of RNA mutations is undesired, especially in gene therapy applications of base editing. In crop breeding, base editing

is currently mostly done by stable genetic transformation. However, BE constructs integrated into the genome can be easily segregated out in the T_1 generation, thus eliminating any RNA off-targets [111].

Many of the strategies used to reduce sgRNA-independent off-target DNA editing also reduce sgRNA-independent off-target RNA editing. Rational design approaches led to three highly accurate CBE variants (BE3$^{W90Y/R126E}$, A3A^{R128A}-BE3, and A3A^{Y130F}-BE3) and one ABE variant (ABE7.10^{F148A}) with very low off-target RNA editing activity, while retaining efficient on-target DNA editing activity [91]. By screening BE3 and ABE mutants, two BE3 variants (R33A, R33A/K34A) and two ABE variants (TadA*; K20A/R21A, V82G) were identified that display substantially reduced RNA editing activities [108,109]. In addition, CBEs containing eA3A (eA3A-BE3), hAID (hAID-BE3), or CDA1 (Target-AID) as editing enzymes also exhibit very low RNA editing activity [109]. Also, TadA* adenosine deaminase variants with reduced sgRNA-independent off-target RNA editing were produced by mutation (V106W) or deletion (ΔR153) of individual amino acid residues [110,112].

In addition to undesired deamination at off-target sites in DNA or RNA, BEs can cause unwanted so-called bystander mutations within the activity window. Bystander mutations occur if a C or A nucleotide other than the targeted one is deaminated, or multiple Cs or As within the editing window undergo editing. While in some application areas (e.g., plant breeding) bystander mutations are not a serious concern, in that such events can simply be discarded, they are more problematic in other areas such as human gene therapy [62]. Consequently, narrowing the width of the editing window has become a major goal of BE optimization efforts. One of two successful strategies has been to develop deaminase variants, in which the interaction between the active site of the deaminase and the spacer nucleotides is altered. For example, changes of amino acid residues involved in catalysis and DNA binding within APOBEC1 (W90Y, R126E, or R132E; BE3 variants YE1, YE2, or YEE) reduced the size of the editing window from approximately five to approximately two nucleotides [84]. Likewise, introduction of a mutation (F148A) into the TadA domain of BE ABE7.10 (initially made to reduce RNA editing activity of the BE) substantially reduced the size of its editing window [91]. Although such deaminase engineering approaches can reduce bystander editing, they often entail trade-offs between editing efficiency and editing precision [63,84]. The second strategy has been to modify the linker between the deaminase domain and the Cas protein by shortening the linker length and/or changing the linker properties (e.g., increasing stiffness) to reduce deaminase access to bystander Cs or As. For example, replacing the canonical XTEN linker in the CBE BE3 by shorter rigid linkers or removing part of the C terminus of the deaminase (CDA1 or A3A), narrows the editing window of the resulting BEs (BE-PAPAP, CDA1Δ variants, and A3AΔ-BE3 variants; Table 1, Key table) to one to two nucleotides while maintaining high on-target editing efficiency [63,64].

Engineering of BEs for base transversions and dual base editing

In canonical CBEs, a UGI domain is included to protect the edited U from excision by UNG. By contrast, replacement of UGI with UNG enhances removal of the edited U from the DNA strand, thus leaving behind an apurinic/apyrimidinic (AP) site that triggers BER. Based on these considerations, a new type of BEs, dubbed glycosylase BEs or CGBEs (C-to-G BEs), was developed by fusing UNG or other BER proteins (e.g., XRCC1) to the C terminus of nCas9. This approach ultimately allowed achievement of C-to-G transversions in mammalian cells and C-to-A transversions in *Escherichia coli* [113–117]. Interestingly, current CGBEs possess an exceptionally narrow editing window, converting C to G preferentially at position C6 (counted from the PAM-distal end of the 20-bp spacer, with the PAM representing positions 21 to 23) [113]. In theory, when coupled

Key table

Table 1. Major base editors (ABEs and CBEs) and their characteristics

Base editor	Architecture	Editing window[a]	Sequence preference	Refs
BE3	rAPOBEC1-XTEN-nCas9-UGI	C_4–C_8	TC	[11]
Target-AID	nCas9-linker-PmCDA1-UGI	C_2–C_5	None	[29,63]
CDA1-BE3	PmCDA1-XTEN-nCas9-UGI	C_1–C_7	None	[63]
CDA1Δ-BE3	PmCDA1Δ-nCas9-UGI	C_3–C_4	None	[63]
YE1-BE3	rAPOBEC1 (W90Y/R126E)-XTEN-nCas9-UGI	C_5–C_7	TC	[84]
YE2-BE3	rAPOBEC1 (W90Y/R132E)-XTEN-nCas9-UGI	C_5–C_6	TC	[84]
YEE-BE3	rAPOBEC1 (W90Y/R126E/R132E)-XTEN-nCas9-UGI	C_5–C_6	TC	[84]
SECURE-BE3	rAPOBEC1 (R33A or R33A/K34A)-XTEN-nCas9-UGI	C_5–C_7	TC	[108]
SaBE3	rAPOBEC1-XTEN-Sa nCas9-UGI	C_3–C_{12}	TC	[84]
dCpf1-BE	rAPOBEC1-XTEN-dCpf1-UGI	C_8–C_{13}	TC	[85]
A3A-BE3	hA3A-XTEN-nCas9-UGI	C_4–C_8	None	[69–71]
eA3A-BE3	hA3A(N57G)-XTEN-nCas9-UGI	C_4–C_8	TC	[69]
A3AΔ-BE3	hA3AΔ-nCas9-UGI	C_5–C_6	None	[64]
BE-PLUS	GCN4(10×)-nCas9 scFv-rAPOBEC1-UGI	C_4–C_{14}	TC	[78]
AID-BE3	hAID-XTEN-nCas9-UGI	C_3–C_8	None	[53]
TAM	dCas9-linker-hAID(P182X)	C_4–C_8	None	[66]
CRISPR-X	dCas9/MS2-linker-hAIDΔ	C_{-50}–C_{50}	None	[65]
A3G-BE3	hA3G-XTEN-nCas9-UGI	C_4–C_8	CC	[53]
eA3G-BE	hA3G-CTD-XTEN-nCas9-2*UGI	C_4–C_8	CC	[73,74]
ABE7.10	TadA-linker-evoTadA-linker-nCas9	A_4–A_7	None	[15]
ABE7.10^{F148A}	TadA ^{F148A}-evoTadA ^{F148A}-linker-nCas9	A_5	None	[91]
dCasMINI-ABE	TadA-linker-evoTadA-linker- dCasMINI	A_3–A_4	None	[86]
CP-ABEs	TadA -linker-evoTadA -linker-CP-nCas9s	A_4–A_{12}	None	[88]

[a]The editing window is given as distance from the protospacer adjacent motif (PAM) where a base editor (BE) can act to induce efficient base changes. Nucleotide positions are numbered relative to the PAM-distal end of the spacer (e.g., the PAM of SpCas9 corresponds to positions 21 to 23), except for Cas12-based BEs (i.e., dCpf1-BE and dCasMINI-ABE) where the first nucleotide downstream of the PAM corresponds to position 1.

with PAM-less Cas9 variants, this feature should enable precise base conversion at any C in the genome. Soon after their initial description, CGBEs have been successfully applied to various organisms, including plants [118] and animals [119]. It is important to note that, although efficient C-to-A conversion can be achieved with CGBEs in *E. coli*, this appears to be far more challenging in eukaryotic cells.

CBEs and ABEs can each only deaminate a single nucleobase type (i.e., C or A). To enable simultaneous conversion of two different bases, new BEs called dual BEs have been developed by fusing both cytidine deaminase and adenosine deaminase domains to nCas9. Indeed, this approach facilitated the simultaneous introduction of both C-to-T and A-to-G mutations into the target sequence [79–82]. By and large, these dual BEs combine the properties of the corresponding CBE and ABE in a single bifunctional protein with respect to both on-target and (DNA and RNA) off-target activities.

More recently, glycosylase BEs and CBEs were also successfully combined to generate a new dual BE type dubbed AGBE. AGBEs are capable of simultaneously introducing four types of base substitutions into a given spacer sequence with a single sgRNA: C to G, C to T, C to A, and A to G [120]. Dual BEs are very useful, if two types of DNA edits within the target sequence are required, because it is usually more straightforward to use a dual BE rather than attempt sequential editing with two conventional BEs [31]. Moreover dual BEs can produce more amino acid substitutions than the individual BEs, thus offering great potential in saturation mutagenesis and directed evolution [81].

Applications of BEs

Current BEs for eukaryotes can induce all possible base transitions (i.e., purine-to-purine and pyrimidine-to-pyrimidine mutations) and two types of transversions (C to G and G to C) without creating DSBs or requiring repair templates. Since point mutations are the main cause of genetic diseases [47], a major application of BEs is the reversion of mutated disease-causing alleles. For example, BE3 can convert the Alzheimer's disease-associated apolipoprotein E allele *APOE4* to the low-risk *APOE3r* variant in immortalized mouse astrocytes and can correct the cancer-associated p53 Y163C mutation by C-to-T editing [11]. The sickle cell disease was reverted in mice with the help of an ABE (ABE8e-NRCH) by converting the pathogenic allele HBB^S into the benign allele HBB^G in hematopoietic stem cells [121]. Another ABE (ABE8.8) was delivered into the liver of cynomolgus monkeys through lipid nanoparticles to introduce a loss-of-function mutation in the gene *PCSK9* (encoding the lipid metabolism-associated serine protease proprotein convertase subtilisin/kexin type 9), resulting in lower cholesterol levels [122]. Recently, this BE-based treatment has received clearance for clinical trial application and plans to initiate a Phase 1 clinical trial in patients in mid-2022 have been announced (https://ir.vervetx.com/news-releases/news-release-details/verve-therapeutics-announces-clearance-first-verve-101-clinical). Additional hereditary diseases that have been treated by BEs in proof-of-concept studies include Duchenne muscular dystrophy [123], albinism [123,124], Hutchinson-Gilford progeria syndrome [125], and several other metabolic diseases [36,126]. These examples show the great potential of BEs for the future cure of genetic disorders.

In contrast to ABE, CBEs can introduce nonsense mutations in genes of interest by altering one of four codons (CGA, CAG, CAA, and, by editing in the other strand, TGG) into stop codons, thereby producing truncated versions of the encoded protein, which often are dysfunctional [127,128]. Similarly, CGBEs can introduce premature termination codons from TCA and TAC triplets [113]. In contrast to CRISPR/Cas9-mediated gene inactivation, where premature termination codons can be introduced in a largely unpredictable manner (by out-of-frame deletions or insertions), BEs can create stop codons at predefined sites within the target locus, thus providing superior precision and predictability.

Besides precise introduction of point mutations in a targeted manner, BEs have also been used for random mutagenesis to create libraries of point mutations in selected genomic areas to generate allelic diversity for subsequent functional screening [65,66]. For example, random mutagenesis of the *PSMB5* gene encoding the 20S proteasome subunit β type-5, the target of the anticancer drug bortezomib (a proteasome inhibitor), allowed the identification of both known and novel mutations that confer resistance to the drug treatment [65].

In addition to medical science, BEs offer great potential in plant breeding. BEs can quickly and efficiently produce allelic variation in crops [45,71,77,129,130], thus facilitating the introduction of known point mutations associated with improved agronomic traits into elite varieties, or the isolation of novel alleles for desired traits. In numerous proof-of-concept applications, BEs have

been shown to contribute to increased crop yields and improved food quality, generate herbicide resistances, and elevate tolerances to biotic and abiotic stresses. For example, to fine-tune the amylose content of rice, a major contributor to eating and cooking quality, the codons for three key amino acid residues of the Waxy gene product were modified by cytosine base editing [131]. Similarly, BE-mediated editing of the gene for the strawberry transcription factor FabZIPs1.1 enhanced sucrose accumulation and improved fruit flavor [132]. To mimic a naturally occurring polymorphism discovered in pea (*Pisum sativum*), a CBE was used to introduce a virus-resistant *eIF4E1* allele into the homologous *Arabidopsis thaliana* gene, thus successfully generating potyvirus-resistant *Arabidopsis* lines [133]. Combination of BEs with an sgRNA library enabled the rapid development of rice lines that display resistance to the herbicide bispyribac sodium (an acetolactate synthase inhibitor) by introducing mutations into the *ALS1* gene [134]. BEs with broader editing windows, and BEs comprising multiple deaminase domains, generally produce libraries of higher complexity with a larger part of the mutation space covered. Cas9 nickases fused to both cytidine and adenosine deaminase domains confer simultaneous C > T and A > G mutations, thus enabling efficient mutagenesis and rapid evolution of herbicide resistance [81] or other useful traits. Further improvements in BE design will likely make such mutagenesis approaches more saturating and less biased, thus greatly facilitating the systematic generation and functional assessment of allelic diversity in plant and animal breeding.

CRISPR-based prime editors

ABEs and CBEs can efficiently induce base transitions, but are not suitable to generate base transversions, or small insertions or deletions. Some reports have shown that BEs can be engineered to also perform conversions from C to G and C to A [114,115,117,135], or deletion of small DNA fragments [136], albeit at relatively low efficiency and with low precision.

Prime editing represents the latest development in genome editing strategies aimed at achieving: (i) more precise DNA changes (avoiding bystander editing); and (ii) more versatile editing that includes all 12 types of point mutations and also the possibility to produce small insertions or deletions without creating DSBs at the target site [137]. Prime editors are comprised of a fusion protein consisting of a Cas9 nickase and a reverse transcriptase (RT) domain (Figure 3), and a prime editing sgRNA (pegRNA). The pegRNA is a guide RNA with an extension at the 3′ end of the scaffold that contains two additional functional domains: a primer-binding sequence (PBS) complementary to the 3′ end of the nicked DNA strand, and an upstream RNA sequence that serves as a template for reverse transcription. Thus, the pegRNA plays two distinct roles in prime editing. First, like the original sgRNA, it guides the prime editor to its target site, where the displaced strand is cut by the Cas9 nickase activity. Second, the PBS anneals with the nicked DNA strand and the 3′OH end of the DNA is extended by reverse transcription of the RNA template containing the sequence to be introduced by prime editing. cDNA synthesis leads to formation of a branched intermediate at the nicked site, consisting of two DNA flaps: the extended and edited DNA (3′ flap) and the original DNA strand (5′ flap). Hybridization of the two flaps to the uncut DNA strand occurs in an equilibrium that then is resolved by the DNA repair machinery (Figure 3). Removal of the unedited 5′ flap results in formation of a DNA heteroduplex consisting of one edited and one unedited strand. The cellular MMR system can subsequently replace the original sequence with the edited sequence.

The first prime editor version (PE1) used the native RT from Moloney murine leukemia virus fused to the C terminus of the Cas9 nickase. Further modifications of the RT (in prime editor PE2) increased the editing efficiency by approximately threefold. To further support the resolution of the heteroduplex structure by the cellular MMR system towards the edited sequence, an additional sgRNA was introduced (in PE3) to generate another nick in the unedited strand. This favors

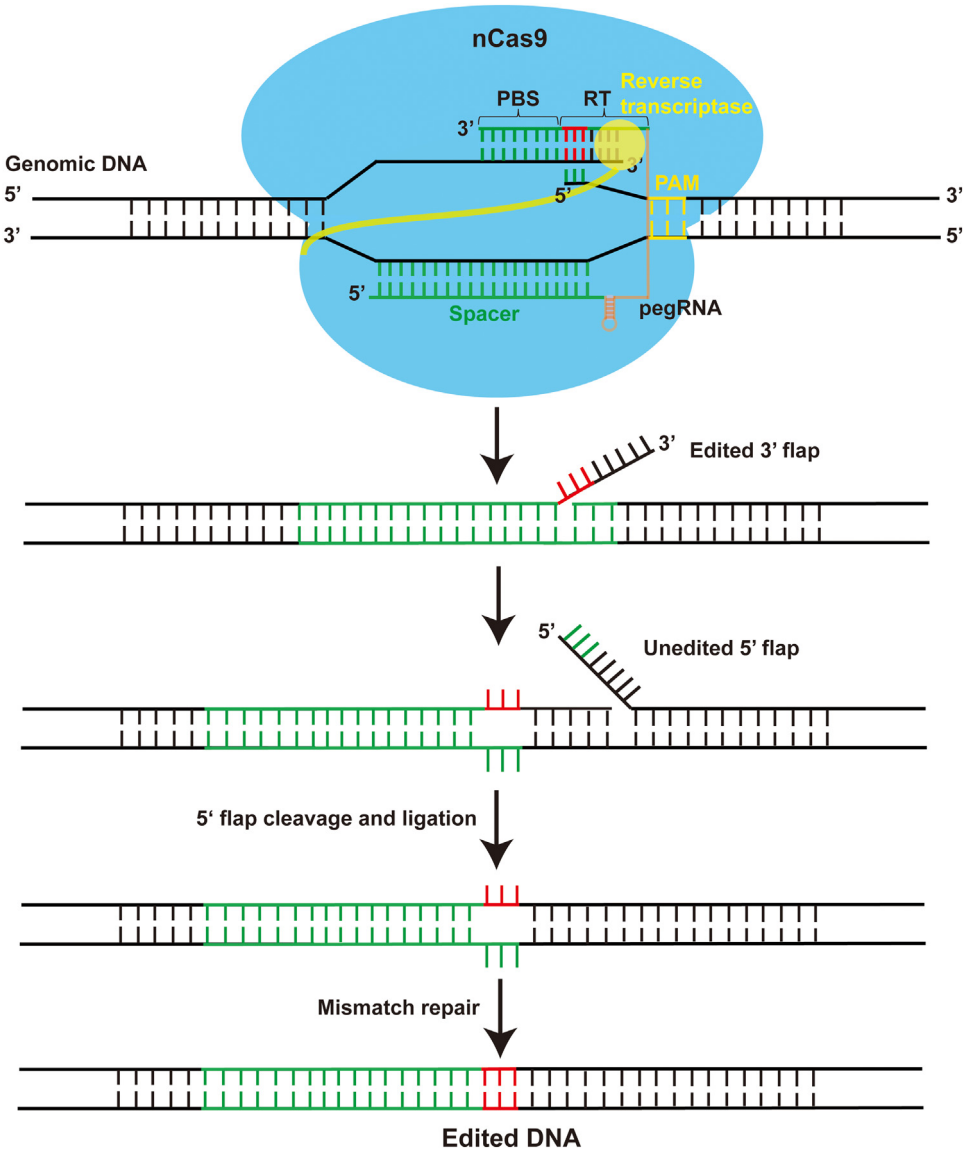

Figure 3. Mechanism of prime editing. Prime editors are typically comprised of a fusion protein of nCas9 with a reverse transcriptase domain and an engineered prime editing single guide RNA (sgRNA; pegRNA). The pegRNA harbors a primer binding site (PBS) and a reverse transcription template (RT) as 3′ extension of the sgRNA scaffold. The pegRNA plays a dual role by: (i) guiding the nCas9 domain to the target site to cut the protospacer adjacent motif (PAM)-containing strand; and (ii) directing the synthesis of an edited DNA strand starting from the 3′ end of the nicked strand and using the RT sequence as template. The 3′ newly synthesized edited DNA strand can displace the previously present 5′ unmodified DNA strand, resulting in a hybridization equilibrium of the uncut strand with two flaps (edited 3′ flap and unedited 5′ flap). The unedited 5′ flap can then be removed by endonuclease digestion, followed by incorporation of the 3′ edited sequence into the target sites to form a heteroduplex DNA consisting of an edited and an unedited strand. By mismatch repair (MMR), the unedited strand can undergo repair using the edited strand as a template, thus fixing the mutations introduced by prime editing.

maintenance of the edited DNA sequence (by enhancing its use as repair template) and increases the editing efficiency by two- to fourfold, but entails the disadvantage of inducing a DSB [138]. To alleviate this problem, the second sgRNA is designed to target the edited sequence (PE3b) so

that the second nick is not introduced into the unedited strand until the newly synthesized strand is fully incorporated into the target site. Since their development, prime editors have been used in various cell types and organisms to achieve precise editing for gene therapy [139–142] and accelerated plant breeding [143–146].

Especially the early version of prime editors suffered from low efficiency in many organisms, including plants [143,146–148]. The pegRNA is believed to be the main component affecting prime editing efficiency [137], although a recent report suggests that also the orientation of the RT relative to the Cas9 nickase affects editing efficiency [144]. The length of both the PBS and RT sequences of the pegRNA (Figure 3) affect editing efficiency and their combined optimization can improve editing efficiency at the target locus [137]. Appropriate software tools have been developed to support optimized pegRNA design [149–151]. The melting temperature (T_m value) of the PBS sequence annealed to the cut DNA strand (Figure 3) also appears to affect prime editing efficiency, with approximately 30°C being reported as optimal for stabilization of the RNA/DNA heteroduplex [152]. Incorporation of structured RNA motifs into the 3′ terminus of pegRNAs results in engineered pegRNAs (epegRNAs) with lower rates of 3′ end degradation, thus also improving the efficiency of prime editors in various human cell types [153]. The cellular DNA MMR system lowers the efficiency of prime editing by correcting the introduced mutations and restoring the wild type allele. This problem can be alleviated by coexpression of an engineered MMR-inhibiting protein (MLH1dn) [154]. While initially, prime editing only allowed insertions of up to ~40 bp and deletions of up to ~80 bp [152], a recent modification of the system uses pairs of pegRNAs and enables generation of larger deletions and insertions [152,155–157].

Although still less efficient than BEs, prime editors can perform genome editing in more versatile ways. Importantly, prime editors can incorporate point mutations into a larger sequence stretch, including multiple mutations that can be up to more than 30 nucleotides away from the cut site in the DNA strand (Figure 3). This feature reduces dependence on PAM availability and provides greater targeting flexibility [137]. Prime editors also largely lack bystander editing, as the edited DNA sequence is determined by the RNA template supplied for reverse transcription. This increased flexibility and editing precision enables prime editors to target >90% of the pathogenic gene variants that are currently listed in the ClinVar database, boding well for future applications of base editing in gene therapy [141,158,159]. Recently, prime editors were successfully tested in different human cell types [137], induced pluripotent stem cells [160], and animal disease models [140,142,161,162]. Prime editors also represent promising tools for precision breeding of crops, as they can accurately incorporate useful genetic variation identified as QTLs into the genomes of elite varieties [145]. Given the considerable progress that has been made with increasing the efficiency of base editing, the simultaneous introduction of multiple alleles for crop improvement may soon be routine.

While prime editing currently represents the most versatile method for precise DNA editing, several challenges remain to be addressed in order to promote its broad applicability. First, despite significant progress, the efficiency of prime editing still varies greatly depending on the type of edits to be made, the surrounding sequence context, the target organism, and other factors, making it difficult to achieve efficient base editing at multiple loci (multiplexing) with prime editors. Second, the large size of the prime editor complex poses a challenge for expression and protein stability in some systems and makes introduction into cells with viral vectors difficult in other systems. Third, while current prime editors possess high target specificity [163,164], whether or not this specificity can be retained in future prime editors with increased activity remains to be seen. If these outstanding issues can be addressed, prime editing could become the base editing method of choice and enable a wide range of new applications.

Genome editing in DNA-containing organelles

The CRISPR/Cas systems currently in use for genome editing in the nucleus are critically dependent on an essential RNA component. Up to now, there is no reliable method available to direct RNAs expressed from nuclear transgenes into one of the two DNA-containing organelles, mitochondria and plastids (chloroplasts). This problem currently precludes the use of the CRISPR/Cas base editing technique for engineering organellar genomes. Ironically, in those cases where direct transformation of organelles (and, therefore, introduction of genes for guide RNAs) is possible, genome editing can be much more easily achieved by **homologous recombination (HR)**, thus making CRISPR/Cas superfluous. Consequently, protein-only genome editing systems, like those based on **transcription activator-like effectors (TALEs)**, currently represent the only useful methods for organellar genome editing (Figure 4), given the ease with which foreign proteins can be targeted to plastids and mitochondria.

DddA-based cytosine base editors (DdCBEs)

The initial breakthrough towards organellar genome editing came with a seminal study showing that the cytidine deaminase DddA from the pathogenic bacterium *Burkholderia cenocepacia* acts on double-stranded DNA and that its catalytically active domain can be split into two nontoxic inert fragments, the activity of which is reconstituted when colocalized on DNA [165]. To create DdCBEs, the two halves of the DddA active domain ($DddA_{tox}$) were fused to the C termini of two TALE scaffolds recognizing adjacent binding sites (Figure 4). Consequently, the two DddA fragments can only come together at the targeted site and act upon the spacer sequence separating the two TALE-binding sequences. For mitochondrial targeting of the proteins, suitable presequences (transit peptide sequences) are added to the N termini of the TALE monomers (Figure 4). A copy of UGI at the C terminus of the DdCBEs increases the mutation frequency by inhibiting uracil-DNA glycosylase, similar to CRISPR-based cytosine BEs. As the deamination reaction initially creates a uracil that is prone to DNA repair, the edit becomes fixed only after incorporation of the complementary adenosine. Introducing vectors expressing mitochondrially targeted DdCBEs into the nuclear genome of human HEK293T cells via **lipofection**, several mitochondrial loci could be edited (i.e., they underwent C-to-T transitions) with an efficiency of up to 49% [165]. Off-target mutation rates in the mitochondrial genome were in the range found for the untreated control (0.024–0.030%) for most DdCBEs, but substantially higher (up to 0.13%) in one case. Like DddA, DdCBEs have a strong substrate preference for cytidines preceded by a thymidine, thus essentially limiting their targeting scope to TC and TCC (TC_n) contexts. DdCBEs come in four permutations, with the $DddA_{tox}$ domain split either after DddA amino acid 1333 or amino acid 1397, and the $DddA_{tox}$ N terminus fused to either the left or the right TALE arm. These alternative designs result in slightly different preferences for cytidines residing in different positions of the target site.

Applying this system to animal mitochondria, it was possible to create mice with base-edited mitochondrial genomes that persisted in a heterochondriomic state and were transmitted through the germline [166,167].

A recent study demonstrated that a similar system based on zinc fingers rather than TALEs also works in the mitochondria of human HEK293T cells [168]. Interestingly, even chimeric systems with one arm of the heterodimer being zinc finger-based and the other one TALE-based were capable of inducing base edits in the mitochondrial DNA.

Very recently, an ABE has been added to the toolbox for organellar base editing [169]. When tested by transfection into human cell lines, the ABE achieved A-to-G conversions at several target sites in the mitochondrial genome with up to 49% efficiency. Upon antibiotic selection, it

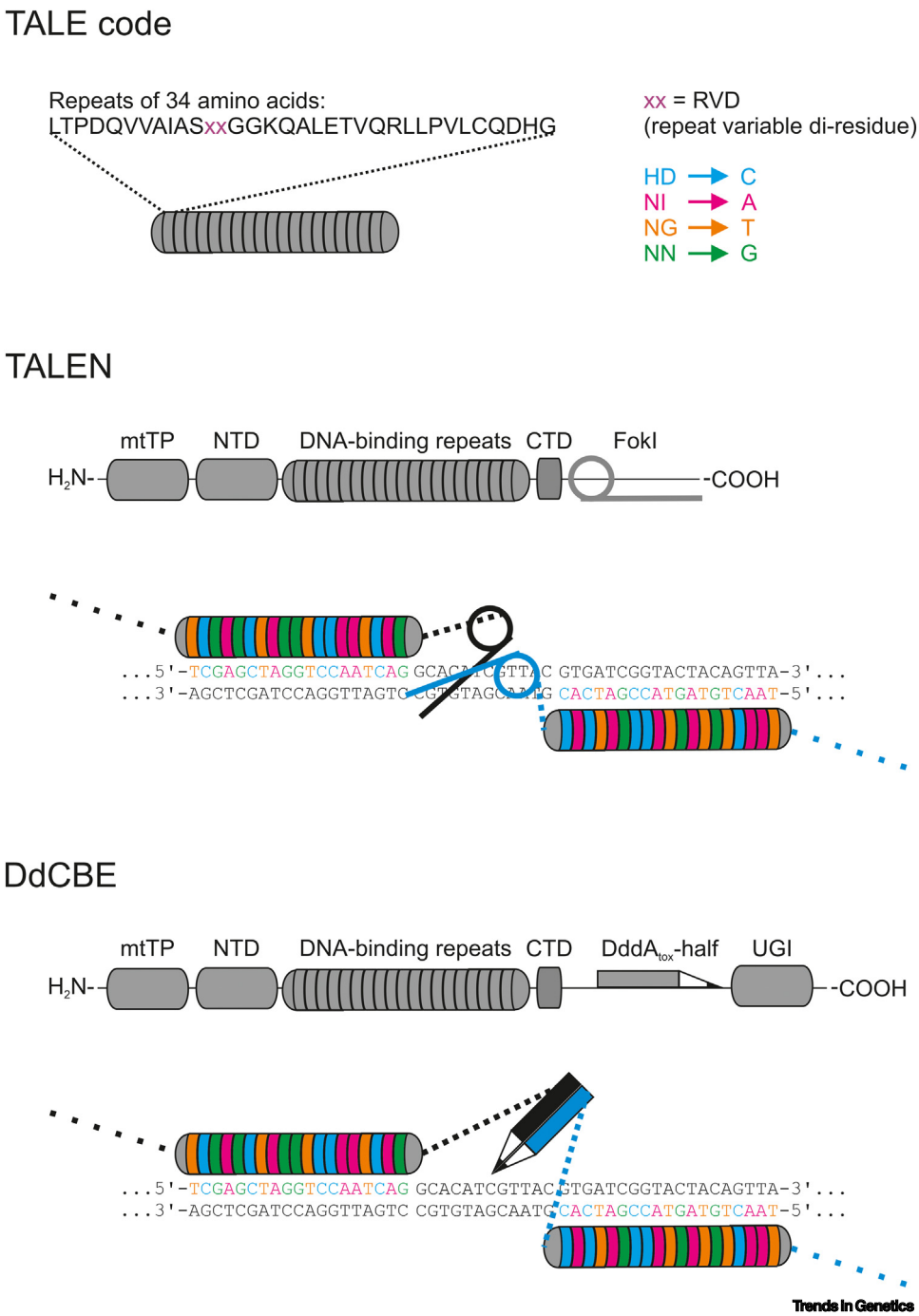

Figure 4. Structure of transcription activator-like effector nucleases (TALENs) and DddA-based cytosine base editors (DdCBEs). Both genome editing tools are based on the transcription activator-like effector (TALE) backbone [185]. A key element is the central array of DNA binding repeats, consisting of units of 34 amino acids each that differ from one another only at positions 12 and 13, the so-called repeat variable di-residues (RVD). Each repeat unit binds one nucleotide in the target DNA and each RVD combination recognizes a specific nucleotide, thus leading to a simple one-to-one TALE code (or RVD-nucleotide code; see list at the right). The number of DNA repeats per TALEN or DdCBE can vary, but normally less than 20 are used. The target sequence can be chosen relatively freely; however, the first nucleotide in the

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

was possible to obtain **homochondriomy** of induced mutations conferring resistance to chloramphenicol. Development of the mitochondrial ABE was facilitated by coupling the engineered TadA8e domain from *E. coli* (that acts only on ssDNA) with a DdCBE. Presumably, the $DddA_{tox}$ domain of DdCBE partially unwinds the double-stranded DNA of the target region, thus making it accessible to the TadA8e adenosine deaminase domain. Consequently, the basal version of this TALE-linked deaminase catalyzes both A-to-G and C-to-T conversions. However, by omitting the UGI domain or using a catalytically dead $DddA_{tox}$ domain (E1347A), a monofunctional ABE could be produced.

Genome editing with DdCBEs in plant organellar genomes

The successful use of DdCBEs for editing organellar genomes in plants was first reported for plastids of *A. thaliana* [170] and plastids and mitochondria of lettuce and rapeseed [171].

Kang *et al*. [171] used PEG-mediated transfection to transiently deliver plasmids encoding DdCBEs to protoplasts of lettuce and rapeseed. The proteins were targeted either to mitochondria via a mitochondrial presequence or to plastids by a chloroplast transit peptide and the strong ubiquitin promoter was used for expression. C-to-T base editing was achieved in the plastid 16S rRNA, in the *psbA* and *psbB* genes (encoding subunits of photosystem II in the thylakoid membrane) and in the mitochondrial *atp6* (encoding a subunit of the ATPase) and *rps14* (encoding a ribosomal protein) genes, with frequencies of up to 30%, as measured 7 days after transfection. Editing efficiency increased further until day 14 and editing events were also observed 4 weeks after transfection in calli derived from transfected protoplasts with frequencies of up to 38%. Using DdCBE-encoding mRNA instead of plasmids for transfection also resulted in editing of the lettuce 16S rRNA locus with a frequency of up to 25%, with indications of reduced off-target editing compared with the use of plasmids, where off-target rates of 1.2% to 4.1% were reported for the sites investigated. A C-to-T editing event in the 16S rRNA target site was predicted to confer resistance to the antibiotic streptomycin. Indeed, when lettuce calli were cultured in the presence of antibiotic selection, the target site was found to be edited to around 98% in antibiotic-resistant calli and to similarly high frequency in shoots derived from streptomycin selection. Notably, the authors provided evidence that, at least in one instance, DdCBEs are able to convert cytosines even in a 5′-AC context. The study also presents a Golden Gate cloning system developed for convenient one-step assembly of plant organellar DdCBEs. All four possible DdCBE combinations (split at 1333 or 1397, DddA N terminus fused to the left or the right TALE arm) were tested.

Nakazato *et al*. [170] went one step further and managed to establish *A. thaliana* plant lines with altered chloroplast genomes that were homoplasmic for the C-to-T transition (i.e., 100% edited). This is particularly important, because cells are highly polyploid with respect to their organellar genomes and the chloroplast genome can easily reach copy numbers of over a thousand per cell [172]. **Homoplasmy** was confirmed in the T_2 generation of the plastid genome-edited lines, after the introduced DdCBE transgene had been eliminated from the nucleus by Mendelian segregation.

binding site must always be a T. The N-terminal repeat unit of the TALE array in the protein binds the 5′ nucleotide of the target site. Parts of the original TALE N- and C-terminal domains (NTD and CTD, respectively) flanking the DNA-binding array are kept in both TALENs and DdCBEs. For application to mitochondrial DNA editing, a mitochondrial presequence (transit peptide, mtTP) is added to the N terminus to target the proteins to mitochondria (for use in plastids, a chloroplast transit peptide would be used instead). For design of TALENs, the C-terminal part of the restriction endonuclease FokI is added at the C terminus (indicated by the half-a-scissor symbol). For DdCBEs, either the N- or the C-terminal part of $DddA_{tox}$ (represented by the half-a-pencil symbol, with $DddA_{tox}$ being the catalytically active part of the double-stranded DNA cytidine editor DddA) is fused to the C terminus. Only when two TALEN or DdCBE monomers bind in tail-to-tail orientation to the target site within a suitable distance (i.e., when their binding sites are separated by a spacer of appropriate length, usually around 14 nucleotides [186]), functional dimers of either FokI (TALEN; indicated by the full scissor) or $DddA_{tox}$ (DdCBE; indicated by the full pencil) are formed and the DNA can either be cut (TALEN) or C-to-T edited (DdCBE) in the spacer region.

In a first step, Nakazato *et al*. established stable transgenic lines with the DdCBE transgenes in the nucleus via ***Agrobacterium*-mediated transformation**. Transgenic T_1 seedlings were then assayed for base editing activity in the target loci (i.e., sequences in the plastid 16S rRNA, *rpoC1* and *psbA* genes) 11 and 23 days after sowing. The majority of seedlings showed on-target editing to at least some extent. Observed editing patterns were complex, ranging from complete to partial or entirely absent editing of individual target nucleotides. In some cases, combinations of fully and partially edited neighboring nucleotides were found in the same plant, indicating that editing occurs within a sequence window and editing efficiency varies with nucleotide position and/or sequence context. The degree of editing also varied between the two time points and, interestingly, both increases and decreases with time were observed. Importantly, 19 out of 51 edited plants showed complete editing of the target nucleotide in all genome copies (homoplasmy). Individual T_2 descendants of several homoplasmic T_1 lines were analyzed for the presence of the mutation induced in the parental generation. The homoplasmic C-to-T conversions were faithfully transmitted into the progeny, displaying the expected uniparentally maternal mode of inheritance that is characteristic of organellar DNA transmission in most seed plants [173]. Stable inheritance occurred also in the absence of the DdCBE transgene in the nucleus, thus excluding *de novo* gene editing in the next generation.

Similar to the Kang *et al*. study [171], all possible orientations for the DdCBEs (split at 1333 or 1397, DddA N terminus at left or right TALE arm; Figure 4) were tested, with the 1397 split appearing to work slightly better than the 1333 split. Although even a 5′-AC site was found to be edited in one instance, not all Cs in the targeted regions were found to be accessible to base editing. Future optimization efforts should involve the systematic sliding of the TALE binding site relative to the targeted nucleotide(s) and length variation and/or structural modification of the spacer separating the DNA-binding domain from the deaminase domain [63,64].

When analyzing the plastid genomes of 17 T_1 plants by next-generation sequencing, some off-target editing events were observed. Off-target edits were found both near the targeted sites and far away in the genome and, in some cases, occurred with up to 100% efficiency. However, no off-target edits were detected in three DdCBE-free homoplasmic T_2 plants, the plastid genomes of which were sequenced. Importantly, no off-target mutations were detected in the mitochondrial genomes of any of the plants. This was important to verify, because it is becoming increasingly clear that protein targeting to plastids and mitochondria is less specific than previously assumed and dual targeting or low-level mistargeting are more frequent than previously appreciated [174,175]. Since mistargeting is likely correlated with expression levels of the deaminase construct from the nuclear genome, it seems advisable to carefully adjust transgene expression levels (by choosing appropriate promoter sequences) and avoid strong overexpression of organellar BE genes. DdCBEs were also shown to enable heritable base editing in the mitochondrial genome of *A. thaliana* [176].

Recently, base editing in the plastid genome has also been achieved in rice [177] by successfully introducing stop codons into the *psaA* gene (encoding a reaction center protein of photosystem I), thus producing nonphotosynthetic mutants. Stable transmission of the edits into the next generation was not reported, presumably because homoplasmic nonphotosynthetic mutants usually do not produce seeds.

TALEN-based editing strategies for organellar genomes

TALENs are genome editing tools that are typically used for disruption of target genes in the nucleus. Upon creation of DSBs by the endonuclease fused to the TALE repeats (Figure 4), the cellular repair machinery often creates small insertions or deletions by NHEJ [178]. TALENs

have also been successfully used to knock out genes in plastids and mitochondria [179–181]. In plant organellar genomes, however, no NHEJ by re-ligation of broken DNA double strands takes place and DNA repair appears to occur predominantly by HR [182]. MMEJ or HR between short to medium size repeats in different regions of the genome can result in elimination of the TALEN target site, thus generating a stable genome that is resistant to TALEN cleavage. The typical products of such recombination events are genomes carrying large deletions that encompass the target site [180]. Unfortunately, size and location of these deletions are nearly unpredictable and often the deleted genomic region will include essential genes, thus preventing the edited genome from reaching the stable homoplasmic state.

Recently, a new method for TALEN-based mutagenesis in plant mitochondria has been developed that is suitable to introduce point mutations into the mitochondrial genome in a targeted fashion [183]. The method, dubbed TALEN-**gene drive** mutagenesis (TALEN-GDM) does not *per se* create mutations, but rather promotes the isolation and purification of pre-existing or induced spontaneous mutations. Plant mitochondria have a very low mutation rate, are highly polyploid [184], and their genomes segregate in a random fashion during cell division. Consequently, it is very difficult to obtain mitochondrial mutations in the first place and isolate homochondriomic mutant plants. Similar to homoplasmy (see earlier), homochondriomy refers to the state in which all mitochondrial genomes of the organism (together forming the chondriome) are identical. Nucleus-encoded TALENs continuously cut their target locus in the mitochondrial genome, which is subsequently restored by HR repair using an intact (uncut) genome copy as repair template [182]. The equilibrium between genome cleavage and HR repair is broken if a mutation occurs in the TALEN binding site that lowers the affinity of the TALEN to the target site (Figure 5). As a result, the mutated genome copies are cut less efficiently than the wild type genomes, which has two effects. First, it likely gives the mutated genomes a maintenance and replication advantage. Second, the mutated genomes serve more often as repair template in HR repair, which essentially will convert all wild type alleles into mutant alleles (Figure 5). Thus, in the presence of the TALENs, the genomes harboring the mutant allele have a direct selective advantage over those containing the wild type allele, representing a gene drive situation. In this way, homochondriomic plants with mutations in the TALEN binding site can be isolated with relative ease (Figure 5) [183].

It was demonstrated that application of the DNA-intercalating agent ethidium bromide further accelerates the appearance and purification of mitochondrial mutations, thus enhancing the efficiency of TALEN-GDM [183]. Whereas BEs that rely on cytidine deaminase or adenosine deaminase fusions enable only certain types of mutations (i.e., C-to-T or A-to-G transitions), any type of point mutation can be recovered with TALEN-GDM. Various transitions and transversions in the mitochondrial *nad9* gene (encoding a subunit of the NADH-ubiquinone oxidoreductase complex of the respiratory chain in the inner mitochondrial membrane, also known as complex I) of tobacco plants have been isolated and purified to homochondriomy [183]. No obvious sequence context bias has been observed and genome resequencing has not detected any off-target mutations outside the TALE-binding site. Given that, to date, no prime editor has been developed for the editing of organellar genomes, the TALEN-GDM method currently provides the most versatile tool for base editing in organelles. It is less site-selective than conventional BEs (in that the entire TALE-binding site is the target of the mutagenesis; Figure 4), but if it is conducted at a sufficiently large scale to be saturating, any desired point mutation within a target sequence of approximately 19 bp can be obtained (Figures 4 and 5). Although so far reported only for mitochondrial genomes, the three basic requirements for TALEN-GDM (presence of a protein import machinery, absence of NHEJ, high activity of HR repair) are also met by plastids and, therefore, the method should be applicable to plastid genomes as well.

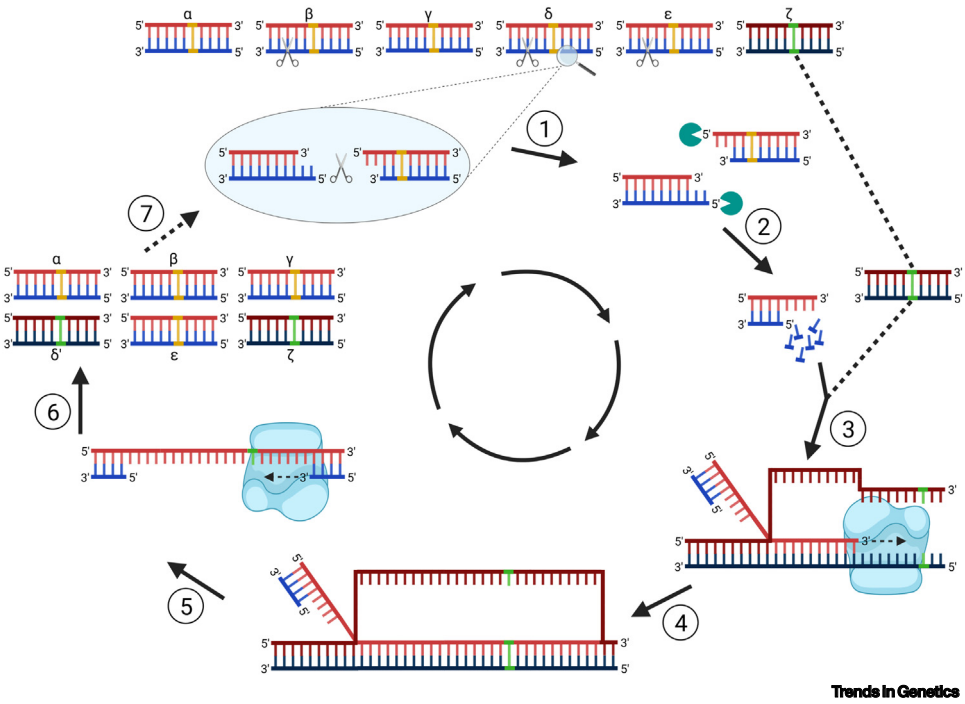

Figure 5. The principle of transcription activator-like effector nuclease (TALEN)-gene-drive mutagenesis (TALEN-GDM). Mitochondria are highly polyploid in that they contain multiple copies of their genome (exemplified by six DNA double strands labeled with Greek letters). If in any of the genomes, a point mutation (green) occurs in the TALEN-binding site, this genome (dark red/dark blue DNA strands) is not cut or is cut at lower frequency than a nonmutated genome (light red/light blue DNA strands, with the nonmutated wild type nucleotide marked in yellow) in the presence of TALEN proteins (scissors) targeting this site. Upon induction of a double-strand break by TALEN activity (1), protruding 3′ ends can be created by exonucleolytic activity (2). The double-strand break is repaired by homologous recombination. The broken strands can invade the homologous sequence in an uncut copy of the mitochondrial genome (3). Repair involves *de novo* DNA synthesis (by a DNA polymerase depicted in light blue) using an intact DNA strand as template (4). If the cut molecule harbors the wild type nucleotide and the intact copy contains the mutated nucleotide, allele conversion takes place (4,5), resulting in an increase in the number of mutated genomes and a decrease in the number of wild type genomes (6; represented as change from genome δ to δ′). Allele conversion can also happen in the other direction, but since genomes with mutations are not cut or are cut less frequently than wild type genomes, the mutant genomes serve, on average, more often as repair template than the wild type genomes. This sets a gene-drive effect in motion (6,7), in that the mutated allele has a selective advantage over the wild type allele directly at the DNA level. Consequently, as soon as a mutation occurs that confers (at least partial) resistance to TALEN cleavage, the mutation will quickly reach the homochondriomic state after a few cycles of genome replication, cleavage, and repair (7).

Concluding remarks

Base editing has created unprecedented new opportunities for high-precision genetic engineering. With the current acceleration of research in this area, and the increasing exploitation of (combinations of) rational design and experimental evolution approaches, we will likely continue to witness the development of new base editing tools at a rapid pace. In addition to improved versions of currently available BEs that offer superior precision and efficiency, entirely new types of BEs will likely become available. Novel types of BEs could rely on new principles of sequence-specific target site recognition and/or comprise enzymatic activities other than nucleoside deaminases and RTs. With the increasing efficiency of base editing methods, the opportunities for multiplexing will also increase. The simultaneous base editing of multiple (unlinked) loci in a genome in one go will make base editing technologies increasingly attractive to synthetic biology and will also enable new applications in experimental evolution and *de novo* domestications of plants and animals [48].

Outstanding questions

How can more compact base editors be designed that are less bulky than the currently used large fusion proteins, but yet edit at high efficiency?

What is the maximum achievable editing accuracy of ABEs and CBEs? Can bystander editing be prevented completely?

Can the DNA repair machinery be tricked into exclusively using the edited DNA strand as repair template?

Can base editing enzymes other than nucleoside deaminases be developed to generate new BEs, including BEs that efficiently induce transversions rather than transitions?

Can sgRNA-like molecules be targeted to organelles to make plastid and mitochondrial genomes accessible to CRISPR/Cas-derived base editors and prime editors?

How can the design principles for base editing reagents (sgRNAs, pegRNAs, TALENs) be simplified, and how can their functionality and efficiency be made more predictable?

How can the efficiency of transient base editing methods (e.g., by transfection of ribonucleoproteins) be improved to avoid the use of transgenes?

Can a consensus regulatory framework and a consensus ethical framework for base editing applications be developed at the international level?

At the same time, the rapid technological progress, the extension of base editing to new cell types and organisms, and the exploration of new areas of application also create new challenges. Applications in human gene therapy raise pressing ethical questions and applications in agricultural biotechnology pose new regulatory issues. The latter concern, for example, the regulatory status of genetically modified organisms that do not contain transgenes, but are derived from a transgenic intermediate (i.e., a genetically encoded BE that subsequently was crossed out) and the regulatory treatment of simple point mutations that were generated by base editing and are indistinguishable from spontaneous mutations or mutations produced by traditional methods of mutagenesis, such as radiation breeding or chemical mutagenesis[i]. It will be of utmost importance to develop widely accepted international standards and guidelines in these areas to: (i) reach consensus on acceptable therapeutic interventions in humans and acceptable genetic manipulations in animals, and (ii) establish rules for the release of base-edited organisms (plants, microbes, and animals) into the environment (see Outstanding questions). This will be essential to reduce barriers to international trade and impediments to the exchange of scientific material and, at the same time, avoid the legal problems that foreseeably would result from a patchwork of laws and regulations.

Acknowledgments

We thank Dennis Kleinschmidt for help with preparing Figure 5. Figure 5 was created with pre-drawn icons from BioRender (https://biorender.com). Research on base editing in the authors' laboratory was supported by the Max Planck Society, a grant from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC-ADG-2014; grant agreement 669982), and a grant from the European Union (Horizon 2020, Newcotiana, 760331-2) to R.B.

Declaration of interests

Three of the authors (R.B., J.T., and D.K.) are inventors on a provisional patent application on high-precision base editors.

Resources

[i]https://www.leopoldina.org/en/publications/detailview/publication/towards-a-scientifically-justified-differentiated-regulation-of-genome-edited-plants-in-the-eu-2019/

References

1. Miller, J.C. *et al.* (2011) A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* 29, 143–148
2. Urnov, F.D. *et al.* (2010) Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* 11, 636–646
3. Jinek, M. *et al.* (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821
4. Jinek, M. *et al.* (2013) RNA-programmed genome editing in human cells. *Elife* 2, e00471
5. Cong, L. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823
6. Mali, P. *et al.* (2013) RNA-guided human genome engineering via Cas9. *Science* 339, 823–826
7. Yeh, C.D. *et al.* (2019) Advances in genome editing through control of DNA repair pathways. *Nat. Cell Biol.* 21, 1468–1478
8. Chang, H.H.Y. *et al.* (2017) Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell Biol.* 18, 495–506
9. Sung, P. and Klein, H. (2006) Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat. Rev. Mol. Cell Biol.* 7, 739–750
10. Molla, K.A. *et al.* (2022) Predictable NHEJ insertion and assessment of HDR editing strategies in plants. *Front. Genome Ed.* 4, 825236
11. Komor, A.C. *et al.* (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424
12. Hess, G.T. *et al.* (2017) Methods and applications of CRISPR-mediated base editing in eukaryotic genomes. *Mol. Cell* 68, 26–43
13. Komor, A.C. *et al.* (2017) CRISPR-based technologies for the manipulation of eukaryotic genomes. *Cell* 168, 20–36
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. 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
16. Gasiunas, G. *et al.* (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 109, E2579–E2586
17. Charpentier, E. *et al.* (2015) Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. *FEMS Microbiol. Rev.* 39, 428–441
18. Chen, H. *et al.* (2014) Cut site selection by the two nuclease domains of the Cas9 RNA-guided endonuclease. *J. Biol. Chem.* 289, 13284–13294
19. Jinek, M. *et al.* (2014) Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* 343, 1247997
20. Nishimasu, H. *et al.* (2014) Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* 156, 935–949
21. Sternberg, S.H. *et al.* (2015) Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature* 527, 110–113
22. Slaymaker, I.M. and Gaudelli, N.M. (2021) Engineering Cas9 for human genome editing. *Curr. Opin. Struct. Biol.* 69, 86–98
23. Kleinstiver, B.P. *et al.* (2015) Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* 523, 481–485
24. Hu, J.H. *et al.* (2018) Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556, 57–63

25. Nishimasu, H. *et al.* (2018) Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* 361, 1259–1262
26. Walton, R.T. *et al.* (2020) Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* 368, 290–296
27. Xu, Z. *et al.* (2021) SpRY greatly expands the genome editing scope in rice with highly flexible PAM recognition. *Genome Biol.* 22, 6
28. Ren, B. *et al.* (2019) Cas9-NG greatly expands the targeting scope of the genome-editing toolkit by recognizing NG and other atypical PAMs in rice. *Mol. Plant* 12, 1015–1026
29. Nishida, K. *et al.* (2016) Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* 353, aaf8729
30. Molla, K.A. and Yang, Y. (2019) CRISPR/Cas-mediated base editing: technical considerations and practical applications. *Trends Biotechnol.* 37, 1121–1142
31. Porto, E.M. *et al.* (2020) Base editing: advances and therapeutic opportunities. *Nat. Rev. Drug Discov.* 19, 839–859
32. Wang, Y. *et al.* (2021) Microbial base editing: a powerful emerging technology for microbial genome engineering. *Trends Biotechnol.* 39, 165–180
33. Yang, B. *et al.* (2019) Development and application of base editors. *CRISPR J.* 2, 91–104
34. Liu, Z. *et al.* (2018) Highly efficient RNA-guided base editing in rabbit. *Nat. Commun.* 9, 2717
35. Zafra, M.P. *et al.* (2018) Optimized base editors enable efficient editing in cells, organoids and mice. *Nat. Biotechnol.* 36, 888–893
36. 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
37. Zhang, Y. *et al.* (2017) Programmable base editing of zebrafish genome using a modified CRISPR-Cas9 system. *Nat. Commun.* 8, 118
38. Li, Q. *et al.* (2018) CRISPR-Cas9-mediated base-editing screening in mice identifies DND1 amino acids that are critical for primordial germ cell development. *Nat. Cell Biol.* 20, 1315–1325
39. Liu, Z. *et al.* (2018) Efficient generation of mouse models of human diseases via ABE- and BE-mediated base editing. *Nat. Commun.* 9, 2338
40. 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
41. Qin, W. *et al.* (2018) Precise A*T to G*C base editing in the zebrafish genome. *BMC Biol.* 16, 139
42. Li, C. *et al.* (2018) Expanded base editing in rice and wheat using a Cas9-adenosine deaminase fusion. *Genome Biol.* 19, 59
43. Liang, P. *et al.* (2017) Effective gene editing by high-fidelity base editor 2 in mouse zygotes. *Protein Cell* 8, 601–611
44. Kang, B.C. *et al.* (2018) Precision genome engineering through adenine base editing in plants. *Nat. Plants* 4, 427–431
45. Hua, K. *et al.* (2018) Precise A.T to G.C base editing in the rice genome. *Mol. Plant* 11, 627–630
46. Molla, K.A. *et al.* (2021) Precise plant genome editing using base editors and prime editors. *Nat. Plants* 7, 1166–1187
47. Landrum, M.J. *et al.* (2016) ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* 44, D862–D868
48. Gao, C. (2021) Genome engineering for crop improvement and future agriculture. *Cell* 184, 1621–1635
49. Zhu, H. *et al.* (2020) Applications of CRISPR-Cas in agriculture and plant biotechnology. *Nat. Rev. Mol. Cell Biol.* 21, 661–677
50. Wood, R.D. (1996) DNA repair in eukaryotes. *Annu. Rev. Biochem.* 65, 135–167
51. Savva, R. *et al.* (1995) The structural basis of specific base-excision repair by uracil-DNA glycosylase. *Nature* 373, 487–493
52. Di Noia, J. and Neuberger, M.S. (2002) Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. *Nature* 419, 43–48
53. 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
54. Wang, L. *et al.* (2017) Enhanced base editing by co-expression of free uracil DNA glycosylase inhibitor. *Cell Res.* 27, 1289–1292
55. Koblan, L.W. *et al.* (2018) Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat. Biotechnol.* 36, 843–846
56. Zhang, X. *et al.* (2020) Increasing the efficiency and targeting range of cytidine base editors through fusion of a single-stranded DNA-binding protein domain. *Nat. Cell Biol.* 22, 740–750
57. Richter, M.F. *et al.* (2020) Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat. Biotechnol.* 38, 883–891
58. Gaudelli, N.M. *et al.* (2020) Directed evolution of adenine base editors with increased activity and therapeutic application. *Nat. Biotechnol.* 38, 892–900
59. Lapinaite, A. *et al.* (2020) DNA capture by a CRISPR-Cas9-guided adenine base editor. *Science* 369, 566–571
60. Li, C. *et al.* (2020) SWISS: multiplexed orthogonal genome editing in plants with a Cas9 nickase and engineered CRISPR RNA scaffolds. *Genome Biol.* 21, 141
61. Jiang, F. and Doudna, J.A. (2017) CRISPR-Cas9 structures and mechanisms. *Annu. Rev. Biophys.* 46, 505–529
62. Anzalone, A.V. *et al.* (2020) Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* 38, 824–844
63. Tan, J. *et al.* (2019) Engineering of high-precision base editors for site-specific single nucleotide replacement. *Nat. Commun.* 10, 439
64. Tan, J. *et al.* (2020) Expanding the genome-targeting scope and the site selectivity of high-precision base editors. *Nat. Commun.* 11, 629
65. Hess, G.T. *et al.* (2016) Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. *Nat. Methods* 13, 1036–1042
66. Ma, Y. *et al.* (2016) Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. *Nat. Methods* 13, 1029–1035
67. Endo, M. *et al.* (2019) Genome editing in plants by engineered CRISPR-Cas9 recognizing NG PAM. *Nat. Plants* 5, 14–17
68. Ren, B. *et al.* (2018) Improved base editor for efficiently inducing genetic variations in rice with CRISPR/Cas9-guided hyperactive hAID mutant. *Mol. Plant* 11, 623–626
69. Gehrke, J.M. *et al.* (2018) An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. *Nat. Biotechnol.* 36, 977–982
70. Wang, X. *et al.* (2018) Efficient base editing in methylated regions with a human APOBEC3A-Cas9 fusion. *Nat. Biotechnol.* 36, 946–949
71. Zong, Y. *et al.* (2018) Efficient C-to-T base editing in plants using a fusion of nCas9 and human APOBEC3A. *Nat. Biotechnol.* 950–953
72. Coelho, M.A. *et al.* (2018) BE-FLARE: a fluorescent reporter of base editing activity reveals editing characteristics of APOBEC3A and APOBEC3B. *BMC Biol.* 16, 150
73. Liu, Z. *et al.* (2020) Precise base editing with CC context-specificity using engineered human APOBEC3G-nCas9 fusions. *BMC Biol.* 18, 111
74. Lee, S. *et al.* (2020) Single C-to-T substitution using engineered APOBEC3G-nCas9 base editors with minimum genome- and transcriptome-wide off-target effects. *Sci. Adv.* 6, eaba1773
75. Cheng, T.L. *et al.* (2019) Expanding C-T base editing toolkit with diversified cytidine deaminases. *Nat. Commun.* 10, 3612
76. Banno, S. *et al.* (2018) Deaminase-mediated multiplex genome editing in *Escherichia coli*. *Nat. Microbiol.* 3, 423–429
77. Shimatani, Z. *et al.* (2017) Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. *Nat. Biotechnol.* 35, 441–443
78. Jiang, W. *et al.* (2018) BE-PLUS: a new base editing tool with broadened editing window and enhanced fidelity. *Cell Res.* 28, 855–861
79. Zhang, X. *et al.* (2020) Dual base editor catalyzes both cytosine and adenine base conversions in human cells. *Nat. Biotechnol.* 38, 856–860
80. Sakata, R.C. *et al.* (2020) Base editors for simultaneous introduction of C-to-T and A-to-G mutations. *Nat. Biotechnol.* 38, 865–869

81. Li, C. *et al.* (2020) Targeted, random mutagenesis of plant genes with dual cytosine and adenine base editors. *Nat. Biotechnol.* 38, 875–882
82. Grunewald, J. *et al.* (2020) A dual-deaminase CRISPR base editor enables concurrent adenine and cytosine editing. *Nat. Biotechnol.* 38, 861–864
83. Xu, R. *et al.* (2021) Development of an efficient plant dual cytosine and adenine editor. *J. Integr. Plant Biol.* 63, 1600–1605
84. 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
85. Li, X. *et al.* (2018) Base editing with a Cpf1-cytidine deaminase fusion. *Nat. Biotechnol.* 36, 324–327
86. Xu, X. *et al.* (2021) Engineered miniature CRISPR-Cas system for mammalian genome regulation and editing. *Mol. Cell* 81, 1–13
87. Oakes, B.L. *et al.* (2019) CRISPR-Cas9 circular permutants as programmable scaffolds for genome modification. *Cell* 176, 254–267
88. Huang, T.P. *et al.* (2019) Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors. *Nat. Biotechnol.* 37, 626–631
89. Liu, Y. *et al.* (2020) A Cas-embedding strategy for minimizing off-target effects of DNA base editors. *Nat. Commun.* 11, 6073
90. Chu, S.H. *et al.* (2021) Rationally designed base editors for precise editing of the sickle cell disease mutation. *CRISPR J.* 4, 169–177
91. Zhou, C. *et al.* (2019) Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature* 571, 275–278
92. Fu, Y. *et al.* (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* 31, 822–826
93. Pattanayak, V. *et al.* (2013) High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* 31, 839–843
94. Hsu, P.D. *et al.* (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31, 827–832
95. Lee, J.K. *et al.* (2018) Directed evolution of CRISPR-Cas9 to increase its specificity. *Nat. Commun.* 9, 3048
96. Xu, W. *et al.* (2019) Multiplex nucleotide editing by high-fidelity Cas9 variants with improved efficiency in rice. *BMC Plant Biol.* 19, 511
97. Kim, D. *et al.* (2019) Genome-wide target specificity of CRISPR RNA-guided adenine base editors. *Nat. Biotechnol.* 37, 430–435
98. Hong, R. *et al.* (2019) Improving the specificity of adenine base editor using high-fidelity Cas9. *bioRxiv* Published online July 23, 2019. https://doi.org/10.1101/712109
99. Fu, Y. *et al.* (2014) Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* 32, 279–284
100. Rose, J.C. *et al.* (2020) Suppression of unwanted CRISPR-Cas9 editing by co-administration of catalytically inactivating truncated guide RNAs. *Nat. Commun.* 11, 2697
101. 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
102. Jang, H.K. *et al.* (2021) High-purity production and precise editing of DNA base editing ribonucleoproteins. *Sci. Adv.* 7, eabg2661
103. Zuo, E. *et al.* (2019) Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science* 364, 289–292
104. Jin, S. *et al.* (2019) Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. *Science* 364, 292–295
105. Doman, J.L. *et al.* (2020) Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors. *Nat. Biotechnol.* 38, 620–628
106. Zuo, E. *et al.* (2020) A rationally engineered cytosine base editor retains high on-target activity while reducing both DNA and RNA off-target effects. *Nat. Methods* 17, 600–604
107. Yu, Y. *et al.* (2020) Cytosine base editors with minimized unguided DNA and RNA off-target events and high on-target activity. *Nat. Commun.* 11, 2052
108. Grunewald, J. *et al.* (2019) Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature* 569, 433–437
109. Grunewald, J. *et al.* (2019) CRISPR DNA base editors with reduced RNA off-target and self-editing activities. *Nat. Biotechnol.* 37, 1041–1048
110. Rees, H.A. *et al.* (2019) Analysis and minimization of cellular RNA editing by DNA adenine base editors. *Sci. Adv.* 5, eaax5717
111. Li, S. *et al.* (2022) A large-scale genome and transcriptome sequencing analysis reveals the mutation landscapes induced by high-activity adenine base editors in plants. *Genome Biol.* 23, 51
112. Li, J. *et al.* (2021) Structure-guided engineering of adenine base editor with minimized RNA off-targeting activity. *Nat. Commun.* 12, 2287
113. Molla, K.A. *et al.* (2020) Base editing landscape extends to perform transversion mutation. *Trends Genet.* 36, 899–901
114. Chen, L. *et al.* (2021) Programmable C:G to G:C genome editing with CRISPR-Cas9-directed base excision repair proteins. *Nat. Commun.* 12, 1384
115. Koblan, L.W. *et al.* (2021) Efficient C*G-to-G*C base editors developed using CRISPRi screens, target-library analysis, and machine learning. *Nat. Biotechnol.* 39, 1414–1425
116. Kurt, I.C. *et al.* (2021) CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nat. Biotechnol.* 39, 41–46
117. Zhao, D. *et al.* (2021) Glycosylase base editors enable C-to-A and C-to-G base changes. *Nat. Biotechnol.* 39, 35–40
118. Sretenovic, S. *et al.* (2021) Exploring C-To-G base editing in rice, tomato, and poplar. *Front. Genome Ed.* 3, 756766
119. Chen, S. *et al.* (2022) Efficient C-to-G base editing with improved target compatibility using engineered deaminase-nCas9 fusions. *CRISPR J.* 5, 2
120. Liang, Y. *et al.* (2022) AGBE: a dual deaminase-mediated base editor by fusing CGBE with ABE for creating a saturated mutant population with multiple editing patterns. *Nucleic Acids Res.* 50, 5384–5399
121. Newby, G.A. *et al.* (2021) Base editing of haematopoietic stem cells rescues sickle cell disease in mice. *Nature* 595, 295–302
122. Musunuru, K. *et al.* (2021) In vivo CRISPR base editing of PCSK9 durably lowers cholesterol in primates. *Nature* 593, 429–434
123. Kim, K. *et al.* (2017) Highly efficient RNA-guided base editing in mouse embryos. *Nat. Biotechnol.* 35, 435–437
124. Cai, H. *et al.* (2019) Efficient gene disruption via base editing induced stop in newt *Pleurodeles waltl*. *Genes (Basel)* 10, 837
125. Koblan, L.W. *et al.* (2021) In vivo base editing rescues Hutchinson-Gilford progeria syndrome in mice. *Nature* 589, 608–614
126. Bose, S.K. *et al.* (2021) In utero adenine base editing corrects multi-organ pathology in a lethal lysosomal storage disease. *Nat. Commun.* 12, 4291
127. 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
128. Kuscu, C. *et al.* (2017) CRISPR-STOP: gene silencing through base-editing-induced nonsense mutations. *Nat. Methods* 14, 710–712
129. Zong, Y. *et al.* (2017) Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. *Nat. Biotechnol.* 35, 438–440
130. Molla, K.A. *et al.* (2020) Single-nucleotide editing for zebra3 and wsl5 phenotypes in rice using CRISPR/Cas9-mediated adenine base editors. *aBIOTECH* 1, 106–118
131. Xu, Y. *et al.* (2021) Fine-tuning the amylose content of rice by precise base editing of the Wx gene. *Plant Biotechnol. J.* 19, 11–13
132. Xing, S. *et al.* (2020) Fine-tuning sugar content in strawberry. *Genome Biol.* 21, 230
133. Bastet, A. *et al.* (2019) Mimicking natural polymorphism in eIF4E by CRISPR-Cas9 base editing is associated with resistance to potyviruses. *Plant Biotechnol. J.* 17, 1736–1750
134. Kuang, Y. *et al.* (2020) Base-editing-mediated artificial evolution of OsALS1 in planta to develop novel herbicide-tolerant rice germplasms. *Mol. Plant* 13, 565–572

135. Yuan, T. *et al.* (2021) Optimization of C-to-G base editors with sequence context preference predictable by machine learning methods. *Nat. Commun.* 12, 4902
136. Wang, S. *et al.* (2020) Precise, predictable multi-nucleotide deletions in rice and wheat using APOBEC-Cas9. *Nat. Biotechnol.* 38, 1460–1465
137. Anzalone, A.V. *et al.* (2019) Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576, 149–157
138. Ran, F.A. *et al.* (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154, 1380–1389
139. Chemello, F. *et al.* (2021) Precise correction of Duchenne muscular dystrophy exon deletion mutations by base and prime editing. *Sci. Adv.* 7, eabg4910
140. Crunkhorn, S. (2021) Prime editing rescues mice from genetic disorders. *Nat. Rev. Drug Discov.* 20, 740
141. Hampton, T. (2020) DNA prime editing: a new CRISPR-based method to correct most disease-causing mutations. *JAMA* 323, 405–406
142. Schene, I.F. *et al.* (2020) Prime editing for functional repair in patient-derived disease models. *Nat. Commun.* 11, 5352
143. Lin, Q. *et al.* (2020) Prime genome editing in rice and wheat. *Nat. Biotechnol.* 38, 582–585
144. Xu, W. *et al.* (2021) A design optimized prime editor with expanded scope and capability in plants. *Nat. Plants* 8, 45–52
145. Xu, R. *et al.* (2021) Identification of herbicide resistance OsACC1 mutations via in planta prime-editing-library screening in rice. *Nat. Plants* 7, 888–892
146. Tang, X. *et al.* (2020) Plant prime editors enable precise gene editing in rice cells. *Mol. Plant* 13, 667–670
147. Li, H. *et al.* (2020) Precise modifications of both exogenous and endogenous genes in rice by prime editing. *Mol. Plant* 13, 671–674
148. Xu, W. *et al.* (2020) Versatile nucleotides substitution in plant using an improved prime editing system. *Mol. Plant* 13, 675–678
149. Anderson, M.V. *et al.* (2021) pegIT - a web-based design tool for prime editing. *Nucleic Acids Res.* 49, W505
150. Hsu, J.Y. *et al.* (2021) PrimeDesign software for rapid and simplified design of prime editing guide RNAs. *Nat. Commun.* 12, 1034
151. Hwang, G.H. *et al.* (2021) PE-Designer and PE-Analyzer: web-based design and analysis tools for CRISPR prime editing. *Nucleic Acids Res.* 49, W499–W504
152. Lin, Q. *et al.* (2021) High-efficiency prime editing with optimized, paired pegRNAs in plants. *Nat. Biotechnol.* 39, 923–927
153. Nelson, J.W. *et al.* (2021) Engineered pegRNAs improve prime editing efficiency. *Nat. Biotechnol.* 40, 402–410
154. Chen, P.J. *et al.* (2021) Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell* 184, 5635–5652
155. Anzalone, A.V. *et al.* (2021) Programmable deletion, replacement, integration and inversion of large DNA sequences with twin prime editing. *Nat. Biotechnol.* 40, 731–740
156. Jiang, T. *et al.* (2021) Deletion and replacement of long genomic sequences using prime editing. *Nat. Biotechnol.* 40, 227–234
157. Choi, J. *et al.* (2021) Precise genomic deletions using paired prime editing. *Nat. Biotechnol.* 40, 218–226
158. Matsoukas, I.G. (2020) Prime editing: genome editing for rare genetic diseases without double-strand breaks or donor DNA. *Front. Genet.* 11, 528
159. Newby, G.A. and Liu, D.R. (2021) In vivo somatic cell base editing and prime editing. *Mol. Ther.* 29, 3107–3124
160. Surun, D. *et al.* (2020) Efficient generation and correction of mutations in human iPS cells utilizing mRNAs of CRISPR base editors and prime editors. *Genes (Basel)* 11, 511
161. Liu, Y. *et al.* (2020) Efficient generation of mouse models with the prime editing system. *Cell Discov.* 6, 27
162. Qian, Y. *et al.* (2021) Efficient and precise generation of Tay-Sachs disease model in rabbit by prime editing system. *Cell Discov.* 7, 50
163. Kim, D.Y. *et al.* (2020) Unbiased investigation of specificities of prime editing systems in human cells. *Nucleic Acids Res.* 48, 10576–10589
164. Jin, S. *et al.* (2021) Genome-wide specificity of prime editors in plants. *Nat. Biotechnol.* 39, 1292–1299
165. Mok, B.Y. *et al.* (2020) A bacterial cytidine deaminase toxin enables CRISPR-free mitochondrial base editing. *Nature* 583, 631–637
166. Guo, J. *et al.* (2022) DdCBE mediates efficient and inheritable modifications in mouse mitochondrial genome. *Mol. Ther. Nucleic Acids* 27, 73–80
167. Lee, H. *et al.* (2021) Mitochondrial DNA editing in mice with DddA-TALE fusion deaminases. *Nat. Commun.* 12, 1190
168. Lim, K. *et al.* (2022) Nuclear and mitochondrial DNA editing in human cells with zinc finger deaminases. *Nat. Commun.* 13, 366
169. Cho, S.I. *et al.* (2022) Targeted A-to-G base editing in human mitochondrial DNA with programmable deaminases. *Cell* 185, 1764–1776
170. Nakazato, I. *et al.* (2021) Targeted base editing in the plastid genome of *Arabidopsis thaliana*. *Nat. Plants* 7, 906–913
171. Kang, B.C. *et al.* (2021) Chloroplast and mitochondrial DNA editing in plants. *Nat. Plants* 7, 899–905
172. Greiner, S. *et al.* (2020) Chloroplast nucleoids are highly dynamic in ploidy, number, and structure during angiosperm leaf development. *Plant J.* 102, 730–746
173. Ruf, S. *et al.* (2007) Determining the transgene containment level provided by chloroplast transformation. *Proc. Natl. Acad. Sci. U. S. A.* 104, 6998–7002
174. Sharma, M. *et al.* (2019) Targeting specificity of nuclear-encoded organelle proteins with a self-assembling split-fluorescent protein toolkit. *J. Cell Sci.* 132, jcs230839
175. Tabatabaei, I. *et al.* (2019) A highly efficient sulfadiazine selection system for the generation of transgenic plants and algae. *Plant Biotechnol. J.* 17, 638–649
176. Nakazato, I. *et al.* (2022) Targeted base editing in the mitochondrial genome of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2121177119
177. Li, R. *et al.* (2021) High-efficiency plastome base editing in rice with TAL cytosine deaminase. *Mol. Plant* 14, 1412–1414
178. Malzahn, A. *et al.* (2017) Plant genome editing with TALEN and CRISPR. *Cell Biosci.* 7, 21
179. Kazama, T. *et al.* (2019) Curing cytoplasmic male sterility via TALEN-mediated mitochondrial genome editing. *Nat. Plants* 5, 722–730
180. Arimura, S.I. *et al.* (2020) Targeted gene disruption of ATP synthases 6-1 and 6-2 in the mitochondrial genome of *Arabidopsis thaliana* by mitoTALENs. *Plant J.* 104, 1459–1471
181. Huang, C.H. *et al.* (2021) Repairing TALEN-mediated double-strand break by microhomology-mediated recombination in tobacco plastids generates abundant subgenomic DNA. *Plant Sci.* 313, 111028
182. Kohl, S. and Bock, R. (2009) Transposition of a bacterial insertion sequence in chloroplasts. *Plant J.* 58, 423–436
183. Forner, J. *et al.* (2022) Targeted introduction of heritable point mutations into the plant mitochondrial genome. *Nat. Plants* 8, 245–256
184. Chevigny, N. *et al.* (2020) DNA repair and the stability of the plant mitochondrial genome. *Int. J. Mol. Sci.* 21, 328
185. Becker, S. and Boch, J. (2021) TALE and TALEN genome editing technologies. *Gene Genome Edit.* 2, 100007
186. Christian, M.L. *et al.* (2012) Targeting G with TAL effectors: a comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. *PLoS One* 7, e45383