Check for updates

Progress in Genome Editing Technology and Its Application in Plants

Kai Zhang[1,2], Nadia Raboanatahiry[1], Bin Zhu[1] and Maoteng Li[1,2]*

[1] Department of Biotechnology, College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan, China, [2] Hubei Collaborative Innovation Center for the Characteristic Resources Exploitation of Dabie Mountains, Huanggang Normal University, Huanggang, China

Genome editing technology (GET) is a versatile approach that has progressed rapidly as a mechanism to alter the genotype and phenotype of organisms. However, conventional genome modification using GET cannot satisfy current demand for high-efficiency and site-directed mutagenesis, retrofitting of artificial nucleases has developed into a new avenue within this field. Based on mechanisms to recognize target genes, newly-developed GETs can generally be subdivided into three cleavage systems, protein-dependent DNA cleavage systems (i.e., zinc-finger nucleases, ZFN, and transcription activator-like effector nucleases, TALEN), RNA-dependent DNA cleavage systems (i.e., clustered regularly interspaced short palindromic repeats-CRISPR associated proteins, CRISPR-Cas9, CRISPR-Cpf1, and CRISPR-C2c1), and RNA-dependent RNA cleavage systems (i.e., RNA interference, RNAi, and CRISPR-C2c2). All these techniques can lead to double-stranded (DSB) or single-stranded breaks (SSB), and result in either random mutations via non-homologous end-joining (NHEJ) or targeted mutation via homologous recombination (HR). Thus, site-directed mutagenesis can be induced via targeted gene knock-out, knock-in, or replacement to modify specific characteristics including morphology-modification, resistance-enhancement, and physiological mechanism-improvement along with plant growth and development. In this paper, an non-comprehensive review on the development of different GETs as applied to plants is presented.

Keywords: Genome editing technology, ZFN, TALEN, CRISPR-Cas9, CRISPR-Cpf1, CRISPR-C2c1, RNAi, CRISPR-C2c2

INTRODUCTION

Genome editing is a conventional method that is often applied to alter the genotype and phenotype of organisms. In general, two methods are typically adopted for gene function analysis. One of these is the traditional forward genetics method (i.e., moving from phenotypic to genetic changes) which enables the identification of new functional genes via T-DNA tag or map-based cloning. This method involves cultivation and screening for special mutants with key traits (Page and Grossniklaus, 2002). The second of the two methods is reverse genetics (i.e., moving from genetic to phenotypic changes), and involves the identification of candidate genes with obvious differential expression. This approach requires that tissues and cells in different environments are identified by gene chip or bioinformatics so that gene function can be analyzed using genetic transformation

technology, including over-expression and knock-out (Takahashi et al., 1994). Conventional genome editing technologies (GETs) including natural mutations via hybridization and induced mutations due to physical factors such as x-rays (Stadler, 1928a,b) and fast neutrons (Li et al., 2001; Wu et al., 2005) or chemical factors such as ethylmethanesulfonate (Wu et al., 2005; Wang et al., 2008), sodium azide (Talamè et al., 2008), and diepoxybutane (Suzuki et al., 2008) offer a variety of new materials for the generation of either autogenous or non-autogenous random variation. Such variants can include point mutations, deletions, translocations, and chromosome losses. In addition, a range of technologies, including protoplast fusion via polyethylene glycol (PEG; Kao and Michayluk, 1974), electronastic stimulus, or sendai virus-induced methods (Salts et al., 1985), have also been widely applied as they are efficient for hybridization, enforcing inheritance from two parents. However, although these techniques are able to alter genome sequences to some degree, they cannot satisfy the requirements of targeted genome modification.

Due to the development of new approaches, current GETs are able to induce double-stranded breaks (DSBs) which either promote random mutations via error-prone non-homologous end-joining (NHEJ; Gorbunova and Levy, 1999) or targeted mutations via error-free homologous recombination (HR; Symington and Gautier, 2011). Thus, sequence deletions, insertions, duplications, replacements, and translocations are frequently observed (Brunet et al., 2009; Garneau et al., 2010; Lee et al., 2010, 2012; Friedland et al., 2013; Xiao et al., 2013; Xie et al., 2014). HR, however, only plays a role in late S/G2 phases, while NHEJ can be applied over all phases of the cell cycle (Sonoda et al., 2006). Thus, although HR is favored for the insertion of an exogenous DNA donor at a DSB for specific genome modifications, it is inefficient compared to NHEJ. Theoretically, NHEJ includes both Ku-dependent and Ku-independent NHEJ pathways (Symington and Gautier, 2011); the former is often referred to as "classical" NHEJ, able to induce several nucleotide insertions or/and deletions with the participation of Ku70/80 proteins, while the latter is capable of inducing a larger number of deletions via microhomology-mediated end-joining (Fattah et al., 2010).

An appropriate method for gene knock-out that was both simple and practical was not available for many years. However, since 1998, RNA interference (RNAi) has been used to provide a rapid, low-cost, and high-throughput technology for the analysis of gene function (Fire et al., 1998). Knockdown using RNAi, however, just makes a DNA sequence diversify, and cannot achieve either permanent or complete gene knock-out (Sachse et al., 2005; Marine et al., 2012). Over the last 10 years, a series of new GETs, including ZFN and TALEN, have become fashionable because they are able to knock-out or modify specific targeted sequences. These GETs are based on chimeric proteins composed of sequence-specific DNA binding domains and non-specific DNA cleavage nucleases (Townsend et al., 2009; Mahfouz et al., 2011; Zhang et al., 2013). Nevertheless, the design of a novel protein that is able to recognize specific DNA is complicated; thus, the adaptive immune system (CRISPR-Cas) have been

developed and widely applied for their powerful genome-editing capabilities.

To date, two classes of CRISPR-Cas systems have been developed that differ in the configuration of their effector modules, which was originally classified into several types on the basis of their Cas compositions (Makarova et al., 2015). Class 1 system possessing multi-subunit crRNA-effector complexes. This class encompasses type I and type III as well as a putative type IV, while the class 2 system is characterized by the presence of an effector complex that consists of a single, large Cas protein. This second class encompasses type II and type V as well as type VI (Shmakov et al., 2015; Wright et al., 2016). To date, more attention has been focused on class 2 CRISPR-Cas systems because of their flexible compositional construction that means they can be harnessed to create powerful genome editing tools. These systems enable simple, specific, and economical technology via a RNA-guided process (Cong et al., 2013; Jiang et al., 2013; Jinek et al., 2013; Mali et al., 2013a; Abudayyeh et al., 2016; Kim et al., 2016). Indeed, of the available class 2 systems, CRISPR-Cas9 from the type II CRISPR-Cas system is a RNA-guided DNA cleavage system that has been developed into a powerful GET with high on-target efficiency (Cong et al., 2013). At the same time, CRISPR-Cpf1 (CRISPR from *Prevotella* and *Francisella 1*), type V-A protein, and CRISPR-C2c1 (class 2 candidate 1), a type V-B protein, have also provided active RNA-guided DNA cleavage systems (Shmakov et al., 2015; Kim et al., 2016), while CRISPR-C2c2 (class 2 candidate 2), type VI-A CRISPR-Cas system, has also shown great potential for the targeting and editing of single-stranded RNA (Abudayyeh et al., 2016).

Very recently, a novel but controversial GET of *Natronobacterium gregoryi* Argonaute (NgAgo) has also been developed that is able to break down the limitation of selected target sequences, the secondary structure of single guided RNA, and the conformation of target DNA. Building on these developments, the aim of this paper is to emphasize an non-comprehensive review on different GETs that have been developed in recent years as well as their incomplete applications to plants.

DEVELOPMENT AND COMPARISONS OF GETS

GETs could be divided into two groups. The members of one group are mediated by protein-guided nuclease, such as ZFN and TALEN, while members of the other are mediated by special RNA/DNA-guided nuclease, including RNA-dependent DNA cleavage systems like CRISPR-Cas9, CRISPR-Cpf1, and CRISPR-C2c1, and RNA-dependent RNA cleavage systems like RNAi and CRISPR-C2c2.
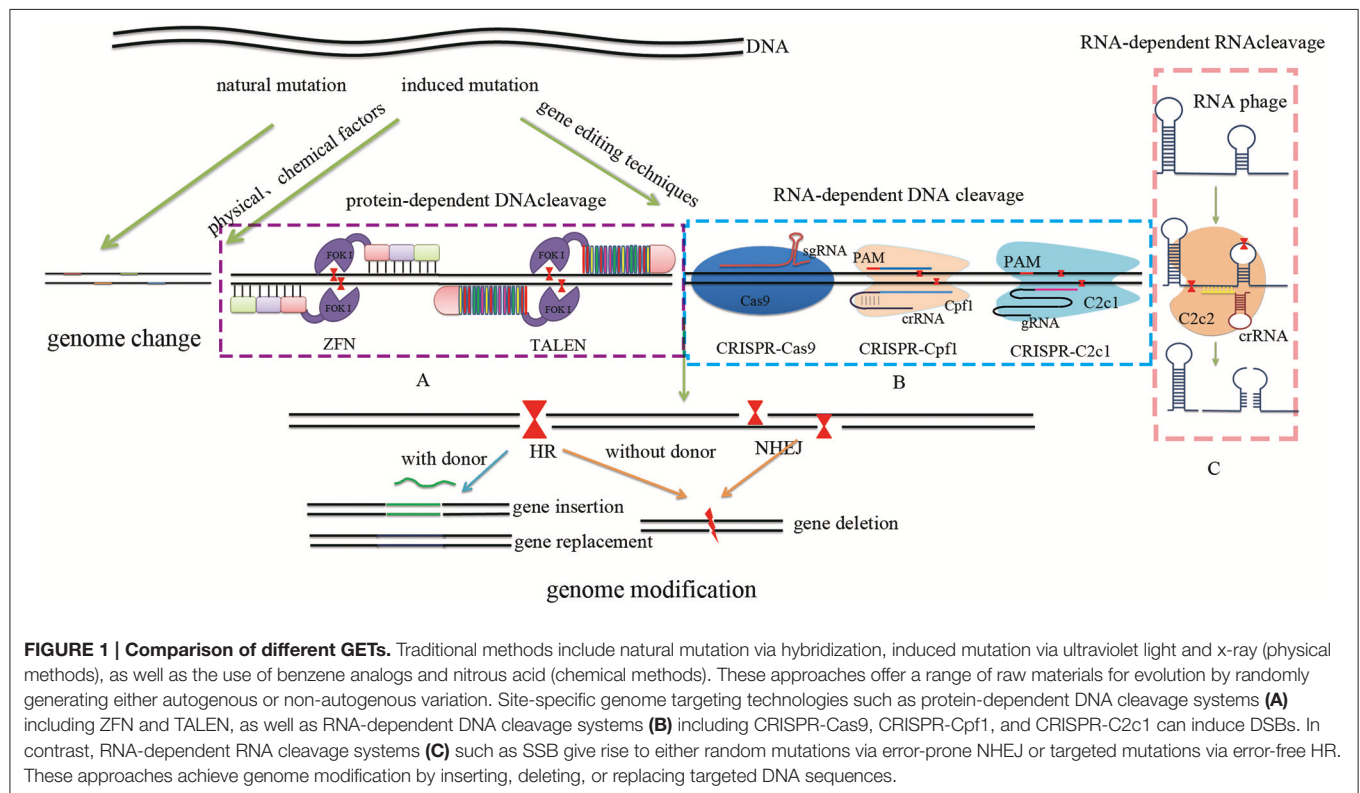
PROTEIN-DEPENDENT DNA CLEAVAGE SYSTEMS

In 1994, Kim and Chandrasegaran developed the first generation of artificial restriction endonucleases by linking the Ultrabithorax

homeodomain of *Drosophila melanogaster* to the cleavage domain of the FokI restriction endonuclease of a *Flavobacterium okeanokoites* (Kim and Chandrasegaran, 1994). Two years later, ZF polypeptides were developed as an alternative to this homeodomain (Kim et al., 1996); the ZF binding domain generally comprises of four-to-six Cys2–His2 arrays which are derived from a transcription factor in human cells. Thus, each ZF is composed of ~30 amino acids arranged in a ββα configuration and mediated by Zn^{2+} (Beerli and Barbas, 2002). Several variable amino acids on the surface of the α-helix at fixed locations are able to bind exclusively to three successive bases in the DNA major groove (Pabo et al., 2001). Therefore, a functional artificial ZFN can be constructed by fusing the ZF binding domain with a restriction enzyme, FokI, extracted from *F. okeanokoites*, which is active as dimers (**Figure 1A**). The key to recognizing specific DNA is the presence of a five-to-seven base pair (bp) spacer between two ZFN target sites that provides a microenvironment for the FokI cleavage domain. The use of ZFNs has proven an optimal strategy for efficient and precise genome editing as these can induce DSBs at a targeted location (Kim et al., 1996; **Figure 1A**). At the same time, however, difficulties inherent to designing ZF binding domains that perfectly match triplet codes as well as the context-dependent nature of interactions with neighboring amino acids have restricted the utilization of this approach (Maeder et al., 2008).

Building on this, a further protein-dependent DNA cleavage platform, TALEN, was listed as a breakthrough of the year in *Science* 2012. This platform, TALE (or AvrBs3/PthA), was discovered in the plant pathogenic bacteria *Xanthomonas*

Hrp-type III secretion (T3S) system (Boch and Bonas, 2010); similarly, linking with FokI enables the introduction of DSBs within site-specific sequences. The TALE comprises three parts, one of which is composed of 13.5–25.5 tandem repeats (i.e., the last repeat is half of a unit in the 3′ terminal) that enables specificity of targeted gene recognition. Each of these repeats usually contains 34 amino acid motifs which are almost identical with the exception of two in positions 12 and 13. This section is referred to as a repeat-variable diresidue (RVD) as it is able to recognize and bind regularly to A, C, T, and G. In contrast, the other parts of TALE include N-terminal and C-terminal sequences (Boch et al., 2009; Moscou and Bogdanove, 2009; **Figure 1A**). As a result of this composition, single base recognition between TALE and DNA-binding repeats is more flexible than ZF proteins which are triplet-limited. At the same time, however, precise recognition by TALE requires a special RVD to identify 5′ methylation C which is sensitive to the target DNA sequence (Boch et al., 2009). In addition, construction of a TALE array presents a technical challenge because of widely identical repeat sequences. More than 1,000 amino acids are required to identify just 20 bases, a process which may lead to an immune response and reduce the efficiency of target recognition (Sander et al., 2011).

Most recently, "base editor"-targeted deaminase technology has been applied to ZF and TALE approaches in order to enhance efficiency and accuracy. Cytosine deaminase, is an enzyme absent in mammalian cells but present in microbes and fungi Building on this observation, Austin and Huber (1993) were the first to clone, sequence, and express *Escherichia coli*. cytosine

FIGURE 1 | Comparison of different GETs. Traditional methods include natural mutation via hybridization, induced mutation via ultraviolet light and x-ray (physical methods), as well as the use of benzene analogs and nitrous acid (chemical methods). These approaches offer a range of raw materials for evolution by randomly generating either autogenous or non-autogenous variation. Site-specific genome targeting technologies such as protein-dependent DNA cleavage systems **(A)** including ZFN and TALEN, as well as RNA-dependent DNA cleavage systems **(B)** including CRISPR-Cas9, CRISPR-Cpf1, and CRISPR-C2c1 can induce DSBs. In contrast, RNA-dependent RNA cleavage systems **(C)** such as SSB give rise to either random mutations via error-prone NHEJ or targeted mutations via error-free HR. These approaches achieve genome modification by inserting, deleting, or replacing targeted DNA sequences.

deaminase and apply it as a gene therapy for the treatment of metastatic colorectal carcinomas (Austin and Huber, 1993). Subsequent research showed that apolipoprotein B editing complexes (APOBECs) and activation-induced deaminase (AID) are also kinds of cytosine deaminase that are able to induce RNA and/or DNA mutations and play a role in the immune system (Muramatsu et al., 1999; Conticello, 2008). These enzymes have been shown to be able to convert cytidines (C) into uracils (U) which themselves are treated as thymine (T). This leads to C = G to T = A conversion in DNA strands unable to induce HR or NHEJ repair mechanisms (Rada and Di Noia, 2004). Yang L. et al. (2016) went further and fused cytosine deaminases to ZF and TALE binding domains, respectively, increasing site-specific C → T conversions by 13% in *E. coli* and 2.5% in human cells. Compared to ZF-APOBECs, ZF-AID possessed a targeted deaminase frequency with highest GFP by correcting ACG to ATG, while this efficiency for TALE-AID fell between ZF-APOBECs and ZF-AID. Almost no off-target conversions were observed when this approach was applied, although several unintended off-target WRC (W = A/T, R = A/G) motifs did result. In addition, research has shown that the cytotoxicity of targeted deaminase ZF-AID is less than is the case for ZFN (Yang L. et al., 2016).

In sum, ZFN and TALEN are composed of DNA-binding protein and the enzyme FokI, both require the elaborate construction of individual DNA-binding proteins for each DNA target site (Zhang et al., 2010; Li et al., 2012).

RNA-DEPENDENT DNA CLEAVAGE SYSTEMS

The CRISPR-Cas series presently provides a robust RNA-dependent DNA cleavage system. Because this approach is based on the adaptive immunity system, widely existing in bacteria and archaea to enable resistance to virus or plasmid invasion, thus showing extreme diversity in Cas protein compositions, genomic loci architecture, and defense mechanisms (Makarova et al., 2011, 2015).

The basic mechanisms of the CRISPR-Cas system-mediated acquired immune system can be roughly divided into three stages (Wiedenheft et al., 2012; Hille and Charpentier, 2016). In the first, adaptation, bacteria and archaea incorporate short sequences from invading genetic elements (i.e., a virus or plasmid) into their genomes. In the second, expression, invading sequences are transcribed and processed into pre-crRNA (precursor CRISPR RNA) under the regulation of leader sequences. When a similar exogenous nucleic acid invades a second time, pre-crRNA sequences can then be further processed into shorter mature crRNA. In the third stage, interference, a tracrRNA-crRNA complex is formed by base pairing and acts to guide the effector protein in recognizing and cleaving the invading foreign DNA sequence homologous to the spacer. Additional research has even elucidated the reason why the CRISPR-Cas system cannot recognize DNA on its own, but just targets sequences; Wiedenheft et al. (2012) argued that this system benefits from the protospacer adjacent motif (PAM) region, and an eight-to-nine bp repetitive sequence in the 5′-end of crRNA that matches the spacer in the CRISPR loci prevent this protein from self-cutting (Wiedenheft et al., 2012; **Figure 2**).

To-date, three RNA-dependent DNA cleavage systems have been reported.

THE CRISPR-CAS9 SYSTEM

The most powerful GET currently, CRISPR-Cas9, was first reported by Jinek et al. (2012). This system is able to recognize a specific site in a target gene with high efficiency, specificity, and flexibility (Jinek et al., 2012; **Figure 1B**). The CRISPR-Cas9 system comprises two main components, one of which is called CRISPR RNA (crRNA), transcribed from interval spacer sequences that correspond to sequences on the phage or plasmid (prospacer). This sequence consists of highly conserved repeats
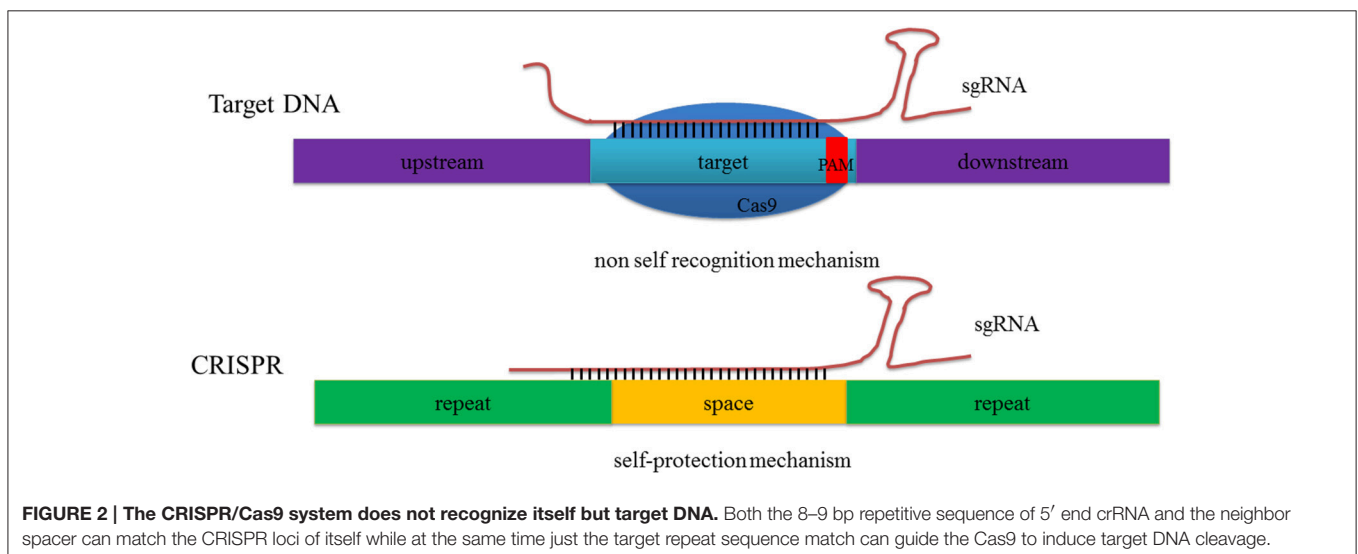

FIGURE 2 | The CRISPR/Cas9 system does not recognize itself but target DNA. Both the 8–9 bp repetitive sequence of 5′ end crRNA and the neighbor spacer can match the CRISPR loci of itself while at the same time just the target repeat sequence match can guide the Cas9 to induce target DNA cleavage.

and markedly different interval spacers that are homologous to the prospacer from the exogenous phage and plasmid (Mojica et al., 2005). The main function of crRNA is to match and recognize target DNA, while another component of this system is Cas protein, close to CRISPR array and responsible for the cleavage of target DNA (Ishino et al., 1987). The numbers of CRISPR as well as its repetitions are highly variable in different species, accounting for the observed diversity in defense functions (Wei et al., 2013). In general, the 5′ or 3′ end of the protospacer has several extended conserved bases, which are called PAMs. Each PAM is between two and five bases in length and is located one-to-four bases away from the protospacer; the composition of each PAM commonly comprises NGG, NRG (R equals G or A) could also be observed in some cases (Hsu et al., 2013; Jiang et al., 2013). Indeed, the one-to-five bp sequence proximal to each PAM in crRNA might determine cleavage specificity and is called the "seed sequence" (Duan et al., 2014; Wu et al., 2014). In contrast, the distal sequences of PAMs are called "non-seed sequences" and may take part in targeting cleavage by triggering conformational changes in Cas9 (Cencic et al., 2014).

The current dominance of the CRISPR-Cas9 approach is the result of several landmark studies that were published at the beginning of 2013 (Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013b). This approach has proved efficient because the Cas9 enzyme has two active centers (RuvC and HNH) that enable strong cutting ability of double-stranded DNA (dsDNA), guided by crRNA, and trans-activating crRNA (tracrRNA). The 5′ end of tracrRNA linked to the 3′ conserved and extended sequence of crRNA may form a hybrid molecule by base pairing, and assists with the formation of a protein-RNA complex by its special space structure with Cas9. This means that this approach is able to further match targeted DNA and achieve DSB by the presence of 20 upstream bases relative to the PAM at the 5′ end. Indeed, the longer the chain on the tracrRNA 3′ end, the better the effect of Cas9 (Hsu et al., 2013). Jinek et al. (2012) creatively merged crRNA with a partial tracrRNA into a chimeric RNA chain and showed that the functions of both can be retained simultaneously (Jinek et al., 2012). As discussed, the chimeric RNA chain dual-tracrRNA:crRNA, which joins crRNA and tracrRNA together into a ring connected by four bases, is called single guide RNA (sgRNA). This is important because only when guided with an appropriate sgRNA, can Cas9 successfully cleave targeted DNA (Jinek et al., 2012). It has even been reported that optimizing the length of an sgRNA sequence may increase on-target efficiency; Fu et al. (2014), for example, reported that a sequence between 17 and 18 nt in length can decrease off-target effects 5,000-times or more without sacrificing on-target efficiency. These authors also speculated that the off-target ratio could be further reduced by increasing GC contents (Fu et al., 2014).

Subsequent to synthesis of Cas9 and gRNA, targeted gene editing can be achieved via delivery into tissues or cells. However, the method utilized to deliver a vector into an animal or plant cell varies a great deal. In animals, for example, a vector containing gRNA and Cas9 can be injected straight into gametes or zygotes. This procedure was pioneered by Cong et al. (2013) who set a procedure for site-specific gene editing in human cells (Cong et al., 2013); this approach is now widely and successfully applied across all kinds of animals. In the case of plants, PEG-mediated transfection to protoplasts (Li et al., 2013; Liang et al., 2014), *Agrobacterium*-mediated transformation to the leaf and embryo (Nekrasov et al., 2013; Upadhyay et al., 2013), and particle bombardment using a gene gun (Miao et al., 2013; Li et al., 2015) are all widely used methods that have led to positive results. Additional research has revealed that codon optimization, as well as the refit of promoters and the number of NLS in Cas9 and sgRNA can affect the efficiency of site-directed delivery (Mikami et al., 2015). For example, the use of either one or two NLS when targeting young seedling albino (YSA) alongside the promoters *Os*U3 and *Os*U6 generates 53.3 vs. 39.6% and 75 vs. 80% mutation rates, respectively (Mikami et al., 2015). The use of codon-optimized Cas9 combined with *Os*Cas9 enables 77.1% on-target efficiency while none were detected when using *At*Cas9 (Mikami et al., 2015). Recently, use of a novel CRISPR/Cas9 vector, pKAMA-ITACHI Red (pKIR), optimized with the ribosomal protein S5A (RPS5A), a Cas9 promoter, enabled both high-efficiency and heritable mutations in *Arabidopsis* (Tsutsui and Higashiyama, 2017). Further, use of the RPS5A optimized-promoter for Cas9 was efficient in 66.7% of albino second leaves in T1 plant cotyledons, while 6.1 and 3.1% were seen, respectively, when WOX2 and 35S optimized-promoters were used to target the PDS3 gene.

When Cas9 and gRNA have been delivered into cells and activated, regulatory mechanisms must be in place to restrain or stop these systems from just carrying out limited knock-out of target genes. To do this, two distinct families of anti-CRISPR associated proteins were identified and used for type I (Pawluk et al., 2016a), while recently a new robust, specific, and genetically encodable "off-switch" for Cas9 activity has been identified (Pawluk et al., 2016b). These authors also identified three completely unrelated sequences of anti-CRISPR families (i.e., $acrIIC1_{Boe}$ and $acrIIC1_{Nme}$, $acrIIC2_{Nme}$, and $acrIIC3_{Nme}$) that can bind directly to the NmeCas9/sgRNA complex and prevent *in vitro* DNA cleavage. Pawluk et al. (2016b) also showed 50% cleavage inhibition when the anti-CRISPR:NmeCas9 ratio was held at 1:1, while complete inhibition was observed at 5:1. A striking decrease in the ability of NmeCas9 to create genome lesions was observed when these anti-CRISPR members were tested in human HEK293T cells (Pawluk et al., 2016b).

Design of Cas9 for Genome Editing

The best-studied Cas endonuclease, Cas9, is a multi-functional protein that has a molecular weight of 160 kDs in type II CRISPR-Cas system. This endonuclease has two structural domains: HNH is responsible for cleaving complementary DNA (cDNA) three nts adjacent to the PAM, while its counterpart, the RuvC-like domain, cleaves strands of non-cDNA in a three-to-eight nt region adjacent to the PAM (Friedland et al., 2013; Xie et al., 2014). Several approaches are available for on-target optimization.

However, an off-target problem will still exist even if this system exhibits outstanding superiority compared to site-specific GET. Thus, researchers have expended large amounts of effort to increase on-target efficiency and decrease the number of off-target events. For example, it has been shown that Cas9
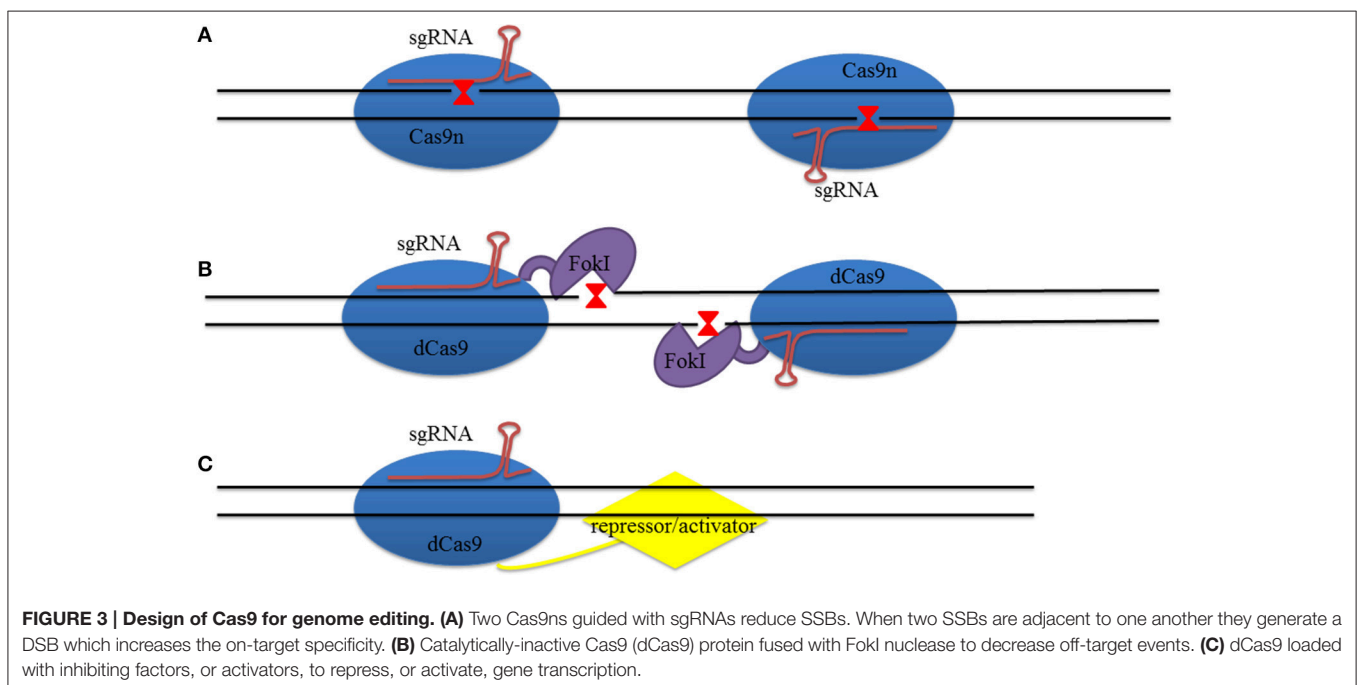
endonuclease can be mutated to Cas9 nickase (Cas9n) that can only induce SSB when RuvC D10A or HNH H840A mutation is introduced (**Figure 3A**). The number of off-target events was decreased more than 50 times when these two Cas9ns were combined with two sgRNA sequences. This is because when these two sequences are present at the same time, the number of mismatches decreases while higher precision HR increases with double Cas9ns (Ran et al., 2013). At the same time, the fCas9 (FokI-dCas9, deactivated Cas9 that has mutations on both the D10A gene of RuvC and the H840A gene of HNH) complex can also be used to reduce the off-target ratio (**Figure 3B**); this approach improved site-specific targeting by fusing FokI with inactivated Cas9 endonuclease, while on-target specificity requires two fCas9 monomers to bind together. The specificity of these modified fCas9s can be up to 140 times higher than use of just wild type Cas9 in human cells (Guilinger et al., 2014). Research has also shown that the use of RNA-guided FokI nuclease (RFN) dimers can also be helpful for reducing the off-target ratio. This ratio was lower when single gRNA guided RFNs were used compared to just Cas9n; in addition, the use of dimeric RFNs can greatly increase on-target efficiency when multiple gRNAs are expressed at the same time and are not restricted by PAMs (Tsai et al., 2014).

Optimized CRISPR 'base editor'-targeted deaminase technology-has also recently been developed, which enables efficient point mutations rather than just random insertions or deletions (indels). This approach also has the advantage that it does not generate DSBs by fusing targeted deaminase with dCas9 guided by sgRNA (Hess et al., 2016; Komor et al., 2016; Ma et al., 2016). Hess et al. (2016) developed this novel approach and performed a mutation in the base region between 12 and 32 bp downstream of the PAM, while Ma et al. (2016) targeted the region between 12 and 16 bp upstream of the PAM (Hess

et al., 2016; Ma et al., 2016). In addition, Hess et al. (2016) were able to successfully alter wild type GFP to brighter EGFP by directed mutation of S65T and F64L residues, obtaining a series of bortezomib-resistant PSMB5 mutants (Hess et al., 2016), while Ma et al. (2016) identified imatinib-resistant mutants by targeting the BCR-ABL gene (Ma et al., 2016). Komor et al. (2016) fused dCas9 with four different cytidine deaminases (i.e., human AID, human APOBEC3G, rat APOBEC1, and lamprey CDA1) and showed that rat APOBEC1-dCas9 had the highest conversion efficiency, as it linked with more than nine amino acids (i.e., BE1, APOBEC1-XTEN-dCas9). In comparison with BE1, the editing efficiency of BE2 (i.e., APOBEC-XTEN-dCas9-UGI) is three-fold higher; both, however, were able to decrease indel rates to <0.1% in human cells. Encouragingly, the base editing efficiency of BE3 (i.e., APOBEC-XTEN-dCas9 (A840H)-UGI) was even higher, an increase of between two and six times compared to BE2 with almost no indel (Komor et al., 2016). These results demonstrate that fusing cytidine deaminase with the CRISPR-Cas system presents a powerful approach for studying and improving gene and protein functions in the future.

THE CRISPR-CPF1 SYSTEM

The cas gene cpf1 was first identified in the *novicida* strain U112 of a subspecies of *Francisella tularensis* (Schunder et al., 2013) and later in other bacteria of this genus as well as in the archaean *Prevotella* (Schunder et al., 2013; Vestergaard et al., 2014; Zetsche et al., 2015). The CRISPR-Cpf1 system, classified as type V-A within class 2, is composed of an ordered cpf1-cas4-cas1-cas2-CRISPR array. Of these, cas1 and cas2 proteins act as adaptation modules involved in spacer DNA acquisition just as in other CRISPR-cas systems (Makarova et al., 2011). In

FIGURE 3 | Design of Cas9 for genome editing. (A) Two Cas9ns guided with sgRNAs reduce SSBs. When two SSBs are adjacent to one another they generate a DSB which increases the on-target specificity. **(B)** Catalytically-inactive Cas9 (dCas9) protein fused with FokI nuclease to decrease off-target events. **(C)** dCas9 loaded with inhibiting factors, or activators, to repress, or activate, gene transcription.

contrast, Cpf1, a characteristic effector module seen in type V-A, is a large protein of about 1,300 amino acids that can be subdivided into a N-terminal helical, a central oligonucleotide-binding domain (OBD), and a C-terminal RuvC domain (Zetsche et al., 2015; Dong et al., 2016). In contrast to the Cas9 system, the N-terminal portion of Cpf1 contains a mixed α/β helical structure, while mostly α-helical recognition lobes are seen in the N-terminal of cas9 (Zetsche et al., 2015). Cpf1 only has a RuvC-like endonuclease domain but lacks a HNH endonuclease domain; thus, the three parts of Cpf1 form a bilobal, triangle-shaped architecture with a large, positively-charged channel in the center (Dong et al., 2016). Research by Dong et al. (2016) has shown that the crystal structure of the full-length LbCpf1 in complex with crRNA is at 2.38Å resolution (Dong et al., 2016). Upstream to the crRNA, PAM loci are located within a 5′-end T-rich region (5′-TTN); Zetsche et al. (2015) showed that of these, the middle T is more important than the first T (Zetsche et al., 2015).

The CRISPR-Cpf1 system is a minimalistic, but functionally adaptive defense system (**Figure 1B**) that was first developed as a robust GET in human cells (Zetsche et al., 2015) and has now been successfully applied in rice (Xu et al., 2016). The constituent Cpf1 is a dual-nuclease, specifically for processing pre-crRNA to mature crRNA, that comprises 19 nt DR fragments followed by a 23–25 nt spacer sequence. In contrast to Cas9, this nuclease does not require tracrRNA or RNaseIII and can also cleave target DNA and introduce DSBs with a crRNA guide consisting of a single stem loop in the DR sequence (Zetsche et al., 2015; Fonfara et al., 2016). Initially, it was thought that Cpf1 forms a homodimer to target DNA cleavage, but later crystal structure studies have revealed that the OBD of this nuclease recognizes crRNA in a highly-distorted conformation. At the same time, the 3′ end of crRNA is directed into the central channel of Cpf1 where extensive intramolecular interactions take place (Dong et al., 2016). Dong et al. (2016) speculated that the octahedral $(Mg(H_2O)6)^{2+}$ ion may play an important role in stabilizing the conformation of crRNA and further enhancing Cpf1 recognition (Zetsche et al., 2015; Dong et al., 2016). The cleavage site is known to occur at, or after, the 23rd nt on the target strand and after the 18th nt on the non-target strand; thus, the length of the 5′ overhang can range between two and five nt and further augment the NHEJ repair mechanism (Zetsche et al., 2015; Kim et al., 2016). The DR portion of mature crRNA should therefore be at least 16 nt in length, and between 17 and 18 nt to maximum cleavage efficiency. Subsequent research has shown that the stem-loop duplex structure of crRNA is more important than the stem loop (Zetsche et al., 2015; Dong et al., 2016); indeed, conserved DR stem-loops are functionally interchangeable between most members of the cpf1 family (Zetsche et al., 2015), while pre-crRNA might be more important than mature crRNA for targeting plant genes. Xu et al. (2016) observed that LbCpf1 transformed with pre-crRNA can enable higher mutation efficiency than the use of mature crRNA in rice (Xu et al., 2016).

Significantly, research has shown that off-target effects can largely be minimized without sacrificing their on-target counterparts when the CRISPR-Cpf1 system is used rather than CRISPR-Cas9 (Kim et al., 2016; Kleinstiver et al., 2016). Results show that Cpf1 tolerates single or double mismatches in the 3′ PAM-distal region, rather than in the 5′ PAM-proximal region (Fonfara et al., 2016; Kim et al., 2016). In addition, Cpf1 preferentially makes deletions as opposed to insertions (Kim et al., 2016). Although simplified Cpf1 enables more convenient GET, challenges remain in on-target cleavage efficiency as this is lower than the better studied Cas9 (Fonfara et al., 2016; Kim et al., 2016).

THE CRISPR-C2C1 SYSTEM

The newly-identified RNA-guided endonuclease, C2c1, belongs to the type V-B CRISPR-Cas system (Shmakov et al., 2015; Yang H. et al., 2016; Liu et al., 2017); **Figure 1B**). The CRISPR-C2c1 system is composed of the ordered c2c1-cas4-cas1-cas2-CRISPR array, similar to the CRISPR-Cpf1 system with the exception of the fact that cas4-cas1 is a fusion protein (Makarova et al., 2011; Shmakov et al., 2015). Indeed, C2c1 is a large protein that comprises between 1,100 and 1,500 amino acids, distinct from any other class 2 effectors (Shmakov et al., 2015). The structure of C2c1 consists of a bi-lobed architecture that resembles a "Crab Claw" consisting of two lobes, an α-helical recognition (REC) lobe that comprises REC1 (helical-I) and REC2 (helical-II), and a NUC lobe that includes wedge (WED, or OBD), RuvC, and Nuc domains (Yang L. et al., 2016; Liu et al., 2017). As a result, C2c1 consists of just a RuvC endonuclease domain and lacks a similar HNH region; RuvC is the most conserved domain of the C2c1 protein, responsible not only for DNA cleavage, but also for sgRNA binding (Liu et al., 2017). In contrast, Nuc is one of the least conserved regions of the C2c1 family and is essential for DNA cleavage activity (Liu et al., 2017). Research by Liu et al. (2017) has shown that the crystal structure of AacC2c1 has 3.1A° resolution when bound to a crRNA-tracrRNA, crRNA is located in the central channel of C2c1, and tracrRNA in positioned in an external surface groove (Liu et al., 2017). This crRNA-tracrRNA duplex can be simplified into chimeric sgRNA, just as is the case in the CRISPR-Cas9 system (Shmakov et al., 2015). Similarly, the 5′-TTN-3′ PAM sequence is analogous to Cpf1, but this system can recognize target DNA in the absence of a PAM-interacting domain (Liu et al., 2017).

This CRISPR-C2c1 system was first shown to be useful for efficient dsDNA cleavage in *Alicyclobacillus acidoterrestris* and subsequently in *Bacillus thermoamylovorans* (Shmakov et al., 2015). C2c1 is a dual-RNA-guided DNA endonuclease rather than being just single-RNA-mediated, as discussed above in the case of Cpf1 (Zetsche et al., 2015; Liu et al., 2017). C2c1 is able to bind sgRNA to form a binary complex and further target DNAs as ternary complexes, which results in staggered breaks (Yang H. et al., 2016; Liu et al., 2017). This cleavage site is known to occur between 14 and 17 bp upstream of the PAM on the target strand and 23 bp upstream of the PAM on the non-target strand; thus, C2c1 results in a staggered DSB with a six-to-eight nt 5′ overhang that facilitates the NHEJ repair mechanism (Liu et al., 2017). It is also noteworthy that because C2c1 is a metal and temperature-dependent endonuclease, highest cleavage activity

is seen in the presence of Mn^{2+} rather than Ni^{2+} or Zn^{2+}, and that temperature between 37 and 60°C are conducive for the most efficient cleavage activity (Shmakov et al., 2015; Yang L. et al., 2016; Liu et al., 2017). In addition, the base-pairing between crRNA repeats and tracrRNA anti-repeats are important for DNA cleavage because single or double mutations directly reduce cleavage efficiency. More importantly, most single-nucleotide mutants of the 18 nt proximal to the PAM sequence in gRNA are able to cease cleavage activity, while the other 2 nts distal to the PAM sequence are only able to reduce activity (Liu et al., 2017). These data provide clear evidence that C2c1 may be an ideal GET for minimal off-target cleavage activity.

RNA-DEPENDENT RNA CLEAVAGE SYSTEMS

The most in-depth RNA-dependent RNA cleavage technique current available is RNAi. This method was first proposed by Fire et al. (1998) because potent gene interference was observed when dsRNA was injected into cells of *Caenorhabditis elegans* (Fire et al., 1998). This approach was later shown to be a credible, efficient, and specific GET for both mammals and plants (Elbashir et al., 2001; Brummelkamp et al., 2002; Travella et al., 2006). Specifically, the mechanism of RNAi mediated gene silencing comprises three stages, the first of which is initiation, where dsRNA is recognized and cleaved by dicer (a RNAseIII-like enzyme) into short segments of 21–23 nt interfering RNA (siRNA). The presence of 2 nt 3′ overhangs and 5′-phosphate termini are essential for the function of siRNA (Zamore et al., 2000). The second stage, effect, involves the combination of siRNA and assembly with an RNA-induced silencing complex (RISC). These RISCs recognize target mRNA, while an anti-sense siRNA strand is paired next to target mRNA via Watson-Crick base pairing. At the same time the sense siRNA strand is released and the entire combined region is cleaved by the RISC (Dzitoyeva et al., 2001). In the final stage, cascade amplification, dsRNA is amplified with templates of target mRNA, primers of siRNA and RNA-dependent RNA polymerase. To induce more siRNA, the effect stage is repeated over and over again so as to silence the target gene. However, RNAi knockdown means that the DNA sequence diversifies, and knock-out neither occurs completely or permanently.

The latest technique to be developed by Abudayyeh et al. (2016) is C2c2 which enables genome editing at the RNA level (Shmakov et al., 2015; Abudayyeh et al., 2016; **Figure 1C**). Indeed, similar to CRISPR-Cas9, C2c2 is a single effector type VI CRISPR-Cas protein that was extracted from the bacterium *Leptotrichia shahii*. However, unlike the types discussed above, C2c2 is specific for a SSB in the target gene and contains two conserved R (N) xxxH motifs, as is typical of the higher eukaryote and prokaryote nucleotide-binding domain (HEPN; Grynberg et al., 2003; Anantharaman et al., 2013). In addition, the HEPN superfamily contains several RNase domains, which can enable further construction of an RNA-dependent cleavage system in bacteria (Kwon et al., 2013). The pre-crRNA in C2c2 can be processed to mature crRNA and further loaded into the C2c2

protein to a mixed α/β secondary structure (Grynberg et al., 2003). Taking into account all these factors, Gao et al. (2016) pointed out that C2c2, as an endoribonuclease, is able to cleave any ssRNA given guidance from a specific 28 nt sequence in crRNA alongside HEPN domains which act as dimers (Gao et al., 2016). Thus, the GET of C2c2 enables a new way to regulate the expression of specific genes and to edit specific effectors at the RNA level.

While C2c2 is a single effector endoRNase which can generate ssRNA cleavage following crRNA, HEPN and crRNA domains are the indispensable elements for cleavage events. The secondary structure of LshC2c2-crRNA is able to bind and target the exposed region of ssRNA as well as having a preference for uracil residues (Abudayyeh et al., 2016). To date, high efficiency cleavage of LshC2c2-crRNA has been achieved using spacer lengths of at least 22 nt as well as the presence of stem-loop structure in the DR. Unlike PAM, however, C2c2 has a bias for A, U, or C in the 3′ protospacer flanking site, while DR is essential for maintenance of the stem loop at an optimum 24 nt length. In contrast, both HEPN domains contain two conserved arginine and histidine residues, which facilitate interactions with LshC2c2 and lead to RNA cleavage.

Another recently developed strategy for transcriptional regulation is the use of nuclease-dCas9. This system loaded with either a repressor or an activator that acts on gene transcription leading to a loss of DNA strand cleavage ability (Cheng et al., 2013; Mali et al., 2013c; Lawhorn et al., 2014; **Figure 3C**). Perez-Pinera et al. (2013), for example, showed that this system is feasible when a VP64 activator is fused to dCas9. This approach led to robust expression of the IL1RN gene in embryonic kidney 293T cells (Perez-Pinera et al., 2013), while Gilbert et al. (2013) demonstrated 93% repression of GFP reporter gene expression when dCas9 was fused with a Krüppel-associated box repressor in human cells (Gilbert et al., 2013). Fusion of EGFP with dCas9 has led to an improved understanding of both the conformation and dynamics of chromosomes at the molecular level (Chen et al., 2013).

THE UTILIZATION OF GETS IN PLANTS

GETs have been widely applied across a range of fields, including for the modification of morphology, enhancement of resistance, and the improvement of physiological mechanisms that are associated with plant growth and development. All of these characteristics can be modified via genome editing involving deletions, insertions, and replacements leading to losses, gains, or changes in the function of target genes.

LOSS-OF-FUNCTION VIA KNOCK-OUT TECHNIQUES

Loss-of-function is the simplest and most widespread mechanism currently applied to gain a better understanding of gene function. To achieve this, knock-out techniques are used as powerful tools to alter specific DNA sequences for the analysis targeted genes. In plants, a number of important agronomical traits

can be improved via site-directed mutagenesis (**Table 1**). For example, the ABA-INSENSITIVE4 (ABI4) gene that regulates the expression of ABA was disrupted in *Arabidopsis* using ZFN. This generated abi4 mutants that were insensitive to ABA and exhibited enhanced stress tolerance to higher glucose concentration (Osakabe et al., 2010). In other work, the locus for mildew resistance (TaMLO), responsible for synthesizing a protein enabling resistance to powdery mildew in plants, was knocked-out by TALEN-induced mutagenesis in allohexaploid bread wheat. This resulted in an allozygote TaMLO gene that imparted higher resistance to powdery mildew disease, further analysis has even shown that homozygous mutations can be transmitted to the next generation (Wang et al., 2014). Following knock-out of the bentazon sensitive lethal (OsBEL) gene, which imparts sensitivity to bentazon and sulfonylurea, 41.2% LbCpf1-induced mutations were identified accompanied with bentazon-resistance (Xu et al., 2016). Finally, RBSDV-resistant plants were obtained under infection pressure in the field when a S7-2-RNAi or S8-RNAi vector was transformed into rice, leading to an overall product improvement (Ahmed et al., 2016).

Some phenotype-related characteristics can also be improved by using fixed-point knock-out techniques. For example, phytoene desaturase gene (PDS) loss-of-function mutations were obtained in T2 bread wheat using RNAi (Travella et al., 2006), Ospds disruption strains induced by CRISSPR/Cas9-meditated mutagenesis have been identified in rice, and albino and dwarf phenotypes are seen as expected (Shan et al., 2013). Similar phenotypes have also been observed in the case of CRISPR-Cas9-induced mutagenesis in *A. thaliana* and *N. tabacum* (Li et al., 2013), while 21.4% Ospds mutations were generated using LbCpf1 system (Xu et al., 2016). A 13.3% proportion of loss-of-function mutants were generated when the chlorophyll a oxygenase 1 gene (cao1) was edited by CRISSPR-Cas9, these plants exhibited a pale green leaf phenotype due to defective synthesis of chlorophyll b (Chl b; Miao et al., 2013). Similarly, a 50% proportion of loss-of-function mutants to the LAZY1 (LA1) gene showed a tiller-spreading phenotype after tillering stage (Miao et al., 2013). In addition, anthocyanin is the flavonoid-derived metabolite which is beneficial to humans because of its strong antioxidant activities (Hou et al., 2004). Zhang et al. (2010) showed that 7% of transparent testa 4 (tt4) gene mutations achieved by application of ZFN showed a deficiency of anthocyanin in the seed coat (Zhang et al., 2010), while the content of phytic acid in maize seeds was successfully reduced by knocking out ZmIPK (an enzyme involved in phytate biosynthesis) using TALEN and CRISPR-Cas9 (Liang et al., 2014). Finally, Soyk et al. (2016) obtained mutations of the self-pruning 5G gene (SP5G), a paralog of florigen that represses flowering activity by applying CRISPR-Cas9 GET. This approach enabled them to successfully broach the limit of day-length sensitivity and create early-yielding varieties (Soyk et al., 2016). In general, these approaches have demonstrated great potential for maize, wheat, and soybean. As for multiple-gene knock-out, it has also been conducted, for example, try cpc etc2 triple mutants that display upwardly-curled leaves as would be expected in *Arabidopsis* (Xing et al., 2014).

GAIN-OF-FUNCTION VIA KNOCK-IN TECHNIQUES

In addition to knock-out in plants, it is also possible to integrate genes with potential applications into the genome to modify specific phenotypic traits (**Table 2**). For example, phosphorus pollution and herbicide resistance are huge current problems; herbicide tolerant and phytate accumulation-mitigated *Zea mays* mutants have been obtained by knocking-out the IPK1 gene and simultaneously introducing a new herbicide-tolerance gene via ZFN (Shukla et al., 2009). Research by Cai et al. (2009) has shown that a phosphinothricin phosphotransferase herbicide resistance gene can be integrated into the endochitinase gene of tobacco via HR using ZFN (Cai et al., 2009). Schiml et al. (2014) demonstrated that the NPTII gene which confers kanamycin resistance can be accurately inserted into the AtADH1 of *A. thaliana* (Schiml et al., 2014), while Li et al. (2015) revealed that hygromycin resistance gene can be conferred into soybean via HR-mediated integration of CRISPR-Cas9 (Li et al., 2015). Besides, Svitashev et al. (2015) showed that bialaphos resistance gene can be integrated into maize via CRISPR-Cas9 (Svitashev et al., 2015). Reporter genes inserted into pre-selected sites will also enable a better understanding of the expression levels of targeted genes accompanied by dynamic changes. For example, Zhang et al. (2013) delivered TALEN plasmids that incorporated yellow fluorescent protein (YFP) and single-strand annealing (SSA) reporters into tobacco protoplasts. This work showed that 14% of targeted YFP insertions were successful (Zhang et al., 2013), while Wang et al. (2014) integrated GFP into TaMLO using TALEN to generate 6.5% green florescence in wheat protoplasts. Further study revealed that these insertions can stabilize inheritance to T1 populations in a Mendelian fashion (Wang et al., 2014). These findings make the introduction of exogenous DNA fragments into plant genes possible.

CHANGE-OF-FUNCTION VIA GENE REPLACEMENT

Gene replacement can be achieved when an extraneous donor is available that possesses a similar terminal with respect to targeted genes. Thus, gene replacement can only be stimulated by DSB when an external DNA donor is available with a homologous terminal to the targeted DNA. When this is the case, this approach can facilitate accurate DNA repair mechanisms and modify genes for a new phenotype or function (**Table 3**).

Weinthal et al. (2013) co-delivered an acceptor DNA molecule (GFP) and a donor DNA molecule (a promoter-less hygromycin B phosphotransferase-encoding gene) flanked by ZFN recognition sequences into *N. tabacum* and *A. thaliana*. The results of this study showed that GFP coding sequences were completely removed along with the recovery of hyp-resistance (Weinthal et al., 2013). In addition, the acetolactate synthase gene (ALS), a major enzyme for herbicides development which can tolerate bispyribac-sodium when two crucial positions of W548L and S627I are mutated, was replaced by a DNA donor of

TABLE 1 | Loss-of-function genome deletion results.

Species	GETs	Target gene	Mutated phenotype	On-targets		Detection method	Transformation technique	Expression system	References
				HR %	NHEJ %				
Brassica napus	RNAi	GPAT4	Abnormal development of several reproductive organs and reduced seed set		N/A	PCR	Agrobacterium-mediated transformation	Stable transgenic	Chen et al., 2014
Solanum lycopersicum	RNAi	Rep/CP	High level resistance against RBSDV		N/A	PCR + Southern hybridization	Agrobacterium-mediated transformation	Stable transgenic	Ammara et al., 2015
Oryza sativa	RNAi	S7-2/S8	Allyl alcohol resistance/lack of anthocyanins in the seed coat		N/A	PCR	Agrobacterium-mediated transformation	Stable transgenic	Ahmed et al., 2016
Arabidopsis thaliana	ZFN	ADH1, TT4	Allyl alcohol resistance/lack of anthocyanins in the seed coat		7/16	PCR + sequencing	Floral-dip transformation	Stable transgenic	Zhang et al., 2010
Arabidopsis thaliana	ZFN	ABI4	ABA and glucose insensitivity		3	Surveyor nuclease assay	Agrobacterium-mediated transformation	Stable transgenic	Osakabe et al., 2010
Glycine max	ZFN	DICER-LIKE (DCL)	Large bulbous internodes		3	PCR method	A.rhizogenes hairy-root transformation	Stable transgenic	Curtin et al., 2011
Solanum lycopersicum	TALEN	PRO	Taller and had lighter green leaves with smoother margins		15	PCR-RE	Agrobacterium-mediated transformation	Stable transgenic	Lor et al., 2014
Triticum aestivum	TALEN	TaMLO	Resistance to powdery mildew		23–38	PCR-RE	Particle bombardment	Stable transgenic	Wang et al., 2014
Oryza sativa	TALEN	OsBADH2	Non-fragrant rice		30	PCR-RE	Particle bombardment	Stable transgenic	Shan et al., 2015
Oryza sativa	TALEN	Os11N3	Resistance to bacterial blight		48	PCR + sequencing	Agrobacterium-mediated transformation	Stable transgenic	Li et al., 2012
Zea mays	TALE	ZmIPK	Reduces phytic acid content in maize seeds		9	PCR/RE + sequencing	Agrobacterium-mediated transformation	Transient expression	Liang et al., 2014
	CRISPR/Cas9				13				
Zea mays	CRISPR/Cas9	LIG1	Leaf with ligule		91	QPCR + sequencing	Biolistic-mediated transformation	Transient expression	Svitashev et al., 2015
		Ms26	Male sterile		77				
		Ms45			100				
		ALS1	Resistance to chlorsulfuron		1.3/0.03				
		ALS2			2.5/2.7				
Glycine max	CRISPR/Cas9	DD20	N.A	4.6	59.3	QPCR assay	Particle bombardment	Stable transgenic	Li et al., 2015
		DD43	N.A	3.8	76		Particle bombardment		

(Continued)

TABLE 1 | Continued

Species	GETs	Target gene	Mutated phenotype	HR %	NHEJ %	Detection method	Transformation technique	Expression system	References
				On-targets					
Arabidopsis thaliana	CRISPR/Cas9	TRY, CPC, ETC2	Upwardly curled leaves		10	T7E1	Floral-dip transformation	Stable transgenic	Xing et al., 2014
Solanum lycopersicum	CRISPR/Cas9	SlAGO7	Compound flat leaves become needle like or wiry		48	PCR + sequencing	A. tumefaciens-mediated transformations	Stable transgenic	Brooks et al., 2014
Oryza sativa	CRISPR/Cas9	OsPDS (SP1)	Albino and dwarf		15	PCR /RE+sequencing	PEG-protoplast transfection	Transient expression	Shan et al., 2013
		OsPDS (SP2)		6.9					
Oryza sativa	CRISPR/Cas9	CAO1	Pale green leaf	13.3	70	PCR + sequencing	Agro-transformation of callus	Stable transgenic	Miao et al., 2013
		LAZY1	Tiller-spreading	50	41.7				
Oryza sativa	CRISPR/Cas9	ROC5	Rice Outermost Cell-specific gene5		26	PCR + RE	Agro-transformation of callus	Stable transgenic	Feng et al., 2013
		SPP	Stromal Processing Peptidase		5				
		YSA	Albino leaf phenotype		48–75				
Arabidopsis thaliana		BRI1	Retarded growth and rolling leaves		26–33	PCR + RE	Agro-transformation by floral dip	Stable transgenic	
		JAZ1	Jasmona-zim-domain protein1		47				
		GAI	Dwarf phenotype		25				
Arabidopsis	CRISPR/Cas9	mir169a	Improved drought resistance		20	PCR + sequencing	Agrobacterium-mediated transformation	Stable transformation	Zhao et al., 2016
Triticum aestivum	CRISPR/Cas9	inox	N.A		17.9–22.3	PCR + sequencing	Agro-transfection of immature embryos	Transient expression	Upadhyay et al., 2013
		pds							
		pds	N.A		18.4–22.3				
Nicotiana benthamiana	CRISPR/Cas9	pds			1.8	PCR + sequencing	Agro-infiltration of leaves	Transient expression	
Triticum aestivum	CRISPR/Cas9	TaMLO	N.A		28.5	PCR + RE	PEG-protoplast transfection	Transient expression	Shan et al., 2013
Arabidopsis thaliana	CRISPR/Cas9	PDS	Leaves with photobleached phenotype	2.70		PCR /RE+sequencing	Agrobacterium leaf infiltration	Transient expression	Li et al., 2013
Nicotiana tabacum				4.8					
Oryza sativa	CRISPR/Cpf1	BEL PDS	Resistence to bentazon and sulfonylurea Albino and dwarf		41.2 21.4	T7E1 assay	Agrobacterium-mediated transformation	Stable expression	Xu et al., 2016

TABLE 2 | Gain-of-function genome insertion results.

Species	GETs	Target gene	Gain-of-function	On-targets		Detection method	Transformation technique	Expression system	References
				HR	NHEJ				
Zea mays	ZFN	IPK1	Herbicide-tolerance	18.6%		Deep sequencing	Whisker-mediated transformation	Stable transgenic	Shukla et al., 2009
Nicotiana tabacum	ZFN	MEL1	Pat herbicide resistance gene cassette	10%		Nested PCR	Agrobacterium transformation	Stable transgenic	Cai et al., 2009
Nicotiana tabacum	TALEN	ALS (SurA, SurB)	YFP		14%	SSA assay	PEG-protoplast transfection	Transient expression	Zhang et al., 2013
Triticum aestivum	TALEN	TaMLO	Protoplasts with GFP		6.5%	PCR-RE	Particle bombardment	Stable transgenic	Wang et al., 2014
	CRISPR/Cas9	TaMLO	His-tag insert to TaMLO-A1/ Myc-tag insert to TaMLO-B1		1/69//3/39	PCR+sequencing	Particle bombardment	Stable transgenic	
Oryza sativa	CRISPR/Cas9	PDS	Single-stranded oligo with a KpnI + EcoRI site	6.8%		PCR /RE+sequencing	PEG-protoplast transfection	Transient expression	Shan et al., 2013
Arabidopsis thaliana	CRISPR/Cas9	ADH1	Kanamycin resistance genes-NPTII	14.6%		PCR +Southern blotting	Agrobacterium-mediated transformation	Stable transgenic	Abudayyeh et al., 2016
Glycine max	CRISPR/Cas9	DD20	Hygromycin resistance	15.5%	57.1%	QPCR assay	Particle bombardment	Stable transgenic	Li et al., 2015
		DD44		2.2%	2.2%				
Zea mays	CRISPR/Cas9	LIG	Resistant to bialaphos	4.1%*	86%*	PCR+sequencing	Particle bombardment	Transient expression	Svitashev et al., 2015
				0	84%		Agrobacterium-mediated transformation	Stable transgenic	

*The frequency obtained when gRNA, Cas9, and donor DNA all in a single vector.

TABLE 3 | Change-of-function genome replacement results.

Species	GETs	Target gene	Changed function	On-targets		Detection method	Transformation technique	Expression system	References
				HR %	NHEJ %				
Arabidopsis	ZFN	QQR-ZFN	GFP to hygromycin-resistant		4.80	PCR + sequencing	A. tumefaciens-mediated transformation	Stable transgenic	Weinthal et al., 2013
Nicotiana tabacum					6.70				
Arabidopsis	ZFN	PPO-loss	Sensitive to the herbicide butafenacil	2		PCR+ Southern blot analysis	Agrobacterium-mediated floral dip transformation	Stable transgenic	de Pater et al., 2013
Nicotiana tabacum	ZFN	SurA, SurB	Resistance to imidazolinone and sulphonylurea herbicides	N.A		Pyrosequencing	Electroporation	Stable transgenic	Townsend et al., 2009
Nicotiana tabacum	TALEN	ALS (SurA, SurB)	Herbicide resistance	4		Protoplast-based single-strand annealing assay	PEG-protoplast transfection	Transient expression	Zhang et al., 2013
Oryza sativa	TALEN	OsALS	Non-herbicides resistance to bispyribac-sodium resistant	6		PCR+sequencing	Ballistic bombardment	Stable transgenic	Li et al., 2016
Pyricularia oryzae	CRISPR/Cas9	SRS2	Bialaphos resistance gene		100.00	PCR analysis	PEG-protoplast transfection	Transient expression	Arazoe et al., 2015
		SDH genescytalone dehydratase	Bialaphos-resistant with white phenotype		36.1–83.6				
Nicotiana tabacum	CRISPR/Cas9	PDS	NbPDS locus to AvrII site	10.7		PCR + sequencing	Agrobacterium leaf infiltration	Transient expression	Li et al., 2013
Arabidopsis	CRISPR/Cas9	AtTFL1	AtTFL1 to eGFP	0.80		PCR + sequencing	Agrobacterium-mediated transformation	Stable transgenic	Zhao et al., 2016
Glycine max	CRISPR/Cas9	Avr4/6	Avr4/6 gene to NPT II gene	13.20		PCR + sequencing	PEG-mediated protoplast transformations	Transient expression	Fang and Tyler, 2016

OsALS mutations with double point mutations through TALEN-based HR in rice. This study showed that up to 6% bispyribac-sodium-resistant rice lines can be generated and are heritable to the T1 generation with normal morphology (Li et al., 2016). In addition, Fang and Tyler (2016) designed a NPT II-based CRISPR-Cas9 cassette flanked by Avr4/6 that comprised three different lengths of 5′/3′ arm sequences (i.e., 250 bp, 500 bp, and 1 kb); the results of this study showed that the longer the attached flanking sequences are, the higher the frequency of induced HR will be. This further corroborates the success of the replacement of Avr4/6 with the NPT II gene (Fang and Tyler, 2016). Zhao et al. (2016) introduced an eGFP expression cassette flanked by *A. thaliana* terminal flower 1 (AtTFL1) left/right homologous arm into a targeted region of this gene via CRISPR-Cas9-meditated replacement; they found that the frequency of 0.8% stably transformed T0 transgenic plants carried eGFP signal and expressed actively in both the leaves and roots of T1 generations (Zhao et al., 2016). Finally, Li et al. (2013) verified HR-induced gene replacement using CRISPR-Cas9 to show that a frequency of 9.0% AvrII incorporation can be achieved when supplying a dsDNA donor that contained a unique AvrII site flanked by left and right homology arm to the NbPDS locus in *N. benthamiana* protoplasts (Li et al., 2013).

CONCLUSIONS

In sum, GETs have great potential for the future analysis of gene function, including for the targeted therapy of human diseases and crop breeding. Stroud et al. (2016), for example, have argued that that GETs coupled with proteomics present powerful tools for studying mitochondrial disease at the cellular level (Stroud et al., 2016). We therefore expect that these technologies will accelerate progress in molecular directional breeding leading to improvements in the quality of cereals

(e.g., rice and wheat) and oil crops (e.g., soybean and oilseed rape). Conventional breeding technologies including the use of chromosome-doubling via the application of colchicine and plasmogamy via PEG are limited to single-species applications. However, given the increasing attention being paid to transgenic food, caution must be exercised in the application of GETs, especially in the case of humans and crops. Taken together, some prerequisites should be followed with regard to both the beneficial and detrimental impacts of GETs on humans. In the first place, potential risks of GETs should be estimated as we move from the laboratory to the field. As unavoidable off-target affects will occur to different degrees, measures such as the analysis and minimization of unfavorable consequences should be taken, recorded, and publicly documented. In addition, the tools used as components of GETs (e.g., ZFN, TALE, and CRISPR) may also have potential cytotoxicity. Secondly, each project that is carried out in this field should encompass a baseline of relative relationships. Third, all satisfactorily-generated transgenic crops or food must be registered and management of this market should be standardized.

AUTHOR CONTRIBUTIONS

KZ carried out the analysis and wrote the manuscript. NR and BZ made helpful suggestions to the manuscript. ML designed, led and coordinated the overall study.

ACKNOWLEDGMENTS

REFERENCES

Abudayyeh, O. O., Gootenberg, J. S., Konermann, S., Joung, J., Slaymaker, I. M., Cox, D. B. T., et al. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* 353:aaf5573. doi: 10.1126/science.aaf5573

Ahmed, M. M. S., Bian, S., Wang, M., Zhao, J., Zhang, B., Liu, Q., et al. (2016). RNAi-mediated resistance to rice black-streaked dwarf virus in transgenic rice. *Transgenic Res.* doi: 10.1007/s11248-016-9999-4. [Epub ahead of print].

Ammara, U. e., Mansoor, S., Saeed, M., Amin, I., Briddon, R. W., and Al-Sadi, A. M. (2015). RNA interference-based resistance in transgenic tomato plants against *Tomato yellow leaf curl virus*-Oman (TYLCV-OM) and its associated beta satellite. *Virol. J.* 12:38. doi: 10.1186/s12985-015-0263-y

Anantharaman, V., Makarova, K. S., Burroughs, A. M., Koonin, E. V., and Aravind, L. (2013). Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. *Biol. Direct* 8:15. doi: 10.1186/1745-6150-8-15

Arazoe, T., Miyoshi, K., Yamato, T., Ogawa, T., Ohsato, S., and Arie, T. (2015). Tailor-made CRISPR/Cas system for highly efficient targeted gene replacement in the rice blast fungus. *Biotechnol. Bioeng.* 112, 2543–2549. doi: 10.1002/bit.25662

Austin, E. A., and Huber, B. E. (1993). A first step in the development of gene therapy for colorectal carcinoma: cloning, sequencing, and expression of *Escherichia coli* cytosine deaminase. *Mol. Pharmacol.* 43, 380–387.

Beerli, R. R., and 1 Barbas, C. R. (2002). Engineering polydactyl zinc-finger transcription factors. *Nat. Biotechnol.* 20, 135–141. doi: 10.1038/nbt0202-135

Boch, J., and Bonas, U. (2010). Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annu. Rev. Phytopathol.* 48, 419–436. doi: 10.1146/annurev-phyto-080508-081936

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., and Kay, S. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326, 1509–1512. doi: 10.1126/science.1178811

Brooks, C., Nekrasov, V., Lippman, Z. B., and Van Eck, J. (2014). Efficient gene editing in tomato in the first generation using the clustered regularly interspaced short palindromic repeats/CRISPR-associated9 system. *Plant Physiol.* 166, 1292–1297. doi: 10.1104/pp.114.247577

Brummelkamp, T. R., Bernards, R., and Agami, R. (2002). A system for stable expression of short interfering RNAs in mammalian cells. *Science* 296, 550–553. doi: 10.1126/science.1068999

Brunet, E., Simsek, D., Tomishima, M., DeKelver, R., Choi, V. M., and Gregory, P. (2009). Chromosomal translocations induced at specified loci in human stem cells. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10620–10625. doi: 10.1073/pnas.0902076106

Cai, C. Q., Doyon, Y., Ainley, W. M., Miller, J. C., Dekelver, R. C., and Moehle, E. A. (2009). Targeted transgene integration in plant cells using designed zinc finger nucleases. *Plant Mol. Biol.* 69, 699–709. doi: 10.1007/s11103-008-9449-7

Cencic, R., Miura, H., Malina, A., Robert, F., Ethier, S., and Schmeing, T. M. (2014). Protospacer adjacent motif (PAM)-distal sequences engage CRISPR

Cas9 DNA target cleavage. *PLoS ONE* 9:e109213. doi: 10.1371/journal.pone.0109213

Chen, B., Gilbert, L. A., Cimini, B. A., Schnitzbauer, J., Zhang, W., and Li, G. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 155, 1479–1491. doi: 10.1016/j.cell.2013.12.001

Chen, X., Chen, G., Truksa, M., Snyder, C. L., Shah, S., and Weselake, R. J. (2014). Glycerol-3-phosphate acyltransferase 4 is essential for the normal development of reproductive organs and the embryo in *Brassica napus*. *J. Exp. Bot.* 65, 4201–4215. doi: 10.1093/jxb/eru199

Cheng, A. W., Wang, H., Yang, H., Shi, L., Katz, Y., and Theunissen, T. W. (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res.* 23, 1163–1171. doi: 10.1038/cr.2013.122

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., and Habib, N. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823. doi: 10.1126/science.1231143

Conticello, S. G. (2008). The AID/APOBEC family of nucleic acid mutators. *Genome Biol.* 9:229. doi: 10.1186/gb-2008-9-6-229

Curtin, S. J., Zhang, F., Sander, J. D., Haun, W. J., Starker, C., and Baltes, N. J. (2011). Targeted mutagenesis of duplicated genes in soybean with zinc-finger nucleases. *Plant Physiol.* 156, 466–473. doi: 10.1104/pp.111.172981

de Pater, S., Pinas, J. E., Hooykaas, P. J., and van der Zaal, B. J. (2013). ZFN-mediated gene targeting of the Arabidopsis protoporphyrinogen oxidase gene through Agrobacterium-mediated floral dip transformation. *Plant Biotechnol. J.* 11, 510–515. doi: 10.1111/pbi.12040

Dong, D., Ren, K., Qiu, X., Zheng, J., Guo, M., and Guan, X. (2016). The crystal structure of Cpf1 in complex with CRISPR RNA. *Nature* 532:522. doi: 10.1038/nature17944

Duan, J., Lu, G., Xie, Z., Lou, M., Luo, J., and Guo, L. (2014). Genome-wide identification of CRISPR/Cas9 off-targets in human genome. *Cell Res.* 24, 1009–1012. doi: 10.1038/cr.2014.87

Dzitoyeva, S., Dimitrijevic, N., and Manev, H. (2001). Intra-abdominal injection of double-stranded RNA into anesthetized adult Drosophila triggers RNA interference in the central nervous system. *Mol Psychiatry* 6, 665–670. doi: 10.1038/sj.mp.4000955

Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 411, 494–498. doi: 10.1038/35078107

Fang, Y., and Tyler, B. M. (2016). Efficient disruption and replacement of an effector gene in the oomycete Phytophthora sojae using CRISPR/Cas9. *Mol. Plant Pathol.* 17, 127–139. doi: 10.1111/mpp.12318

Fattah, F., Lee, E. H., Weisensel, N., Wang, Y., Lichter, N., and Hendrickson, E. A. (2010). Ku regulates the non-homologous end joining pathway choice of DNA double-strand break repair in human somatic cells. *PLoS Genet.* 6:e1000855. doi: 10.1371/journal.pgen.1000855

Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D. L., and Wei, P. (2013). Efficient genome editing in plants using a CRISPR/Cas system. *Cell Res.* 23, 1229–1232. doi: 10.1038/cr.2013.114

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806–811. doi: 10.1038/35888

Fonfara, I., Richter, H., Bratovic, M., Le Rhun, A., and Charpentier, E. (2016). The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* 532, 517. doi: 10.1038/nature17945

Friedland, A. E., Tzur, Y. B., Esvelt, K. M., Colaiácovo, M. P., Church, G. M., and Calarco, J. A. (2013). Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat. Methods* 10, 741–743. doi: 10.1038/nmeth.2532

Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M., and Joung, J. K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* 32, 279–284. doi: 10.1038/nbt.2808

Gao, F., Shen, X. Z., Jiang, F., Wu, Y., and Han, C. (2016). DNA-guided genome editing using the *Natronobacterium gregoryi* Argonaute. *Nat. Biotechnol.* 34, 768–773. doi: 10.1038/nbt.3547

Garneau, J. E., Dupuis, M. È., Villion, M., Romero, D. A., Barrangou, R., and Boyaval, P. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71. doi: 10.1038/nature09523

Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., and Torres, S. E. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154, 442–451. doi: 10.1016/j.cell.2013.06.044

Gorbunova, V. V., and Levy, A. A. (1999). How plants make ends meet: DNA double-strand break repair. *Trends Plant Sci.* 4, 263–269. doi: 10.1016/S1360-1385(99)01430-2

Grynberg, M., Erlandsen, H., and Godzik, A. (2003). HEPN: a common domain in bacterial drug resistance and human neurodegenerative proteins. *Trends Biochem. Sci.* 28, 224–226. doi: 10.1016/S0968-0004(03)00060-4

Guilinger, J. P., Thompson, D. B., and Liu, D. R. (2014). Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol.* 32, 577–582. doi: 10.1038/nbt.2909

Hess, G. T., Frésard, L., Han, K., Lee, C. H., Li, A., Cimprich, K. A., et al. (2016). Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. *Nat. Methods.* 13, 1036–1042. doi: 10.1038/nmeth.4038

Hille, F., and Charpentier, E. (2016). CRISPR-Cas: biology, mechanisms and relevance. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 371:20150496. doi: 10.1098/rstb.2015.0496

Hou, D. X., Kai, K., Li, J. J., Lin, S., Terahara, N., and Wakamatsu, M. (2004). Anthocyanidins inhibit activator protein 1 activity and cell transformation: structure-activity relationship and molecular mechanisms. *Carcinogenesis* 25, 29–36. doi: 10.1093/carcin/bgg184

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., and Agarwala, V. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31, 827–832. doi: 10.1038/nbt.2647

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* 169, 5429–5433. doi: 10.1128/jb.169.12.5429-5433.1987

Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L. A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* 31, 233–239. doi: 10.1038/nbt.2508

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi: 10.1126/science.1225829

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013). RNA-programmed genome editing in human cells. *eLife* 2:e471. doi: 10.7554/eLife.00471

Kao, K. N., and Michayluk, M. R. (1974). A method for high-frequency intergeneric fusion of plant protoplasts. *Planta* 115, 355–367. doi: 10.1007/BF00388618

Kim, D., Kim, J., Hur, J. K., Been, K. W., Yoon, S., and Kim, J. (2016). Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat. Biotechnol.* 34:863. doi: 10.1038/nbt.3609

Kim, Y. G., and Chandrasegaran, S. (1994). Chimeric restriction endonuclease. *Proc. Natl. Acad. Sci. U.S.A.* 91, 883–887. doi: 10.1073/pnas.91.3.883

Kim, Y. G., Cha, J., and Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. U.S.A.* 93, 1156–1160. doi: 10.1073/pnas.93.3.1156

Kleinstiver, B. P., Tsai, S. Q., Prew, M. S., Nguyen, N. T., Welch, M. M., and Lopez, J. M. (2016). Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat. Biotechnol.* 34:869. doi: 10.1038/nbt.3620

Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533:420. doi: 10.1038/nature17946

Kwon, D. N., Lee, K., Kang, M. J., Choi, Y. J., Park, C., and Whyte, J. J. (2013). Production of biallelic CMP-Neu5Ac hydroxylase knock-out pigs. *Sci Rep* 3:1981. doi: 10.1038/srep01981

Lawhorn, I. E., Ferreira, J. P., and Wang, C. L. (2014). Evaluation of sgRNA target sites for CRISPR-mediated repression of TP53. *PLoS ONE* 9:e113232. doi: 10.1371/journal.pone.0113232

Lee, H. J., Kim, E., and Kim, J. S. (2010). Targeted chromosomal deletions in human cells using zinc finger nucleases. *Genome Res.* 20, 81–89. doi: 10.1101/gr.099747.109

Lee, H. J., Kweon, J., Kim, E., Kim, S., and Kim, J. S. (2012). Targeted chromosomal duplications and inversions in the human genome using zinc finger nucleases. *Genome Res.* 22, 539–548. doi: 10.1101/gr.129635.111

Li, J. F., Norville, J. E., Aach, J., McCormack, M., Zhang, D., and Bush, J. (2013). Multiplex and homologous recombination-mediated genome editing

in Arabidopsis and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.* 31, 688–691. doi: 10.1038/nbt.2654

Li, T., Liu, B., Chen, C. Y., and Yang, B. (2016). TALEN-mediated homologous recombination produces site-directed DNA base change and herbicide-resistant rice. *J. Genet. Genomics* 43, 297–305. doi: 10.1016/j.jgg.2016.03.005

Li, T., Liu, B., Spalding, M. H., Weeks, D. P., and Yang, B. (2012). High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat. Biotechnol.* 30, 390–392. doi: 10.1038/nbt.2199

Li, X., Song, Y., Century, K., Straight, S., Ronald, P., and Dong, X. (2001). A fast neutron deletion mutagenesis-based reverse genetics system for plants. *Plant J.* 27, 235–242. doi: 10.1046/j.1365-313x.2001.01084.x

Li, Z., Liu, Z. B., Xing, A., Moon, B. P., Koellhoffer, J. P., and Huang, L. (2015). Cas9-guide RNA directed genome editing in soybean. *Plant Physiol.* 169:960. doi: 10.1104/pp.15.00783

Liang, Z., Zhang, K., Chen, K., and Gao, C. (2014). Targeted mutagenesis in *Zea mays* using TALENs and the CRISPR/Cas system. *J. Genet. Genomics* 41, 63–68. doi: 10.1016/j.jgg.2013.12.001

Liu, L., Chen, P., Wang, M., Li, X., Wang, J., Yin, M., et al. (2017). C2c1-sgRNA complex structure reveals RNA-guided DNA cleavage mechanism. *Mol. Cell.* 65, 310–322. doi: 10.1016/j.molcel.2016.11.040

Lor, V. S., Starker, C. G., Voytas, D. F., Weiss, D., and Olszewski, N. E. (2014). Targeted mutagenesis of the tomato PROCERA gene using transcription activator-like effector nucleases. *Plant Physiol.* 166, 1288–1291. doi: 10.1104/pp.114.247593

Ma, Y., Zhang, J., Yin, W., Zhang, Z., Song, Y., and Chang, X. (2016). Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. *Nat. Methods.* 13, 1029–1035. doi: 10.1038/nmeth.4027

Maeder, M. L., Thibodeau-Beganny, S., Osiak, A., Wright, D. A., Anthony, R. M., and Eichtinger, M. (2008). Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol. Cell* 31, 294–301. doi: 10.1016/j.molcel.2008.06.016

Mahfouz, M. M., Li, L., Shamimuzzaman, M., Wibowo, A., Fang, X., and Zhu, J. K. (2011). *De novo*-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2623–2628. doi: 10.1073/pnas.1019533108

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J. J., Charpentier, E., and Horvath, P. (2011). Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* 9, 467–477. doi: 10.1038/nrmicro2577

Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., and Saunders, S. J. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13, 722–736. doi: 10.1038/nrmicro3569

Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., and Kosuri, S. (2013c). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* 31, 833–838. doi: 10.1038/nbt.2675

Mali, P., Esvelt, K. M., and Church, G. M. (2013a). Cas9 as a versatile tool for engineering biology. *Nat. Methods* 10, 957–963. doi: 10.1038/nmeth.2649

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., and DiCarlo, J. E. (2013b). RNA-guided human genome engineering via Cas9. *Science* 339, 823–826. doi: 10.1126/science.1232033

Marine, S., Bahl, A., Ferrer, M., and Buehler, E. (2012). Common seed analysis to identify off-target effects in siRNA screens. *J. Biomol. Screen.* 17, 370–378. doi: 10.1177/1087057111427348

Miao, J., Guo, D., Zhang, J., Huang, Q., Qin, G., and Zhang, X. (2013). Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Res.* 23, 1233–1236. doi: 10.1038/cr.2013.123

Mikami, M., Toki, S., and Endo, M. (2015). Comparison of CRISPR/Cas9 expression constructs for efficient targeted mutagenesis in rice. *Plant Mol. Biol.* 88, 561–572. doi: 10.1007/s11103-015-0342-x

Mojica, F. J., Díez-Villasenor, C., García-Martinez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* 60, 174–182. doi: 10.1007/s00239-004-0046-3

Moscou, M. J., and Bogdanove, A. J. (2009). A simple cipher governs DNA recognition by TAL effectors. *Science* 326:1501. doi: 10.1126/science.1178817

Muramatsu, M., Sankaranand, V. S., Anant, S., Sugai, M., Kinoshita, K., and Davidson, N. O. (1999). Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase

family in germinal center B cells. *J. Biol. Chem.* 274, 18470–18476. doi: 10.1074/jbc.274.26.18470

Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J. D., and Kamoun, S. (2013). Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* 31, 691–693. doi: 10.1038/nbt.2655

Osakabe, K., Osakabe, Y., and Toki, S. (2010). Site-directed mutagenesis in Arabidopsis using custom-designed zinc finger nucleases. *Proc. Natl. Acad. Sci. U.S.A.* 107, 12034–12039. doi: 10.1073/pnas.1000234107

Pabo, C. O., Peisach, E., and Grant, R. A. (2001). Design and selection of novel Cys2His2 zinc finger proteins. *Annu. Rev. Biochem.* 70, 313–340. doi: 10.1146/annurev.biochem.70.1.313

Page, D. R., and Grossniklaus, U. (2002). The art and design of genetic screens: *Arabidopsis thaliana*. *Nat. Rev. Genet.* 3, 124–136. doi: 10.1038/nrg730

Pawluk, A., Amrani, N., Zhang, Y., Garcia, B., Hidalgo-Reyes, Y., and Lee, J. (2016b). Naturally occurring off-switches for CRISPR-Cas9. *Cell* 167, 1829–1838. doi: 10.1016/j.cell.2016.11.017

Pawluk, A., Staals, R. H. J., Taylor, C., Watson, B. N. J., Saha, S., Fineran, P. C., et al. (2016a). Inactivation of CRISPR-Cas systems by anti-CRISPR proteins in diverse bacterial species. *Nat. Microbiol.* 1:16085. doi: 10.1038/nmicrobiol.2016.85

Perez-Pinera, P., Kocak, D. D., Vockley, C. M., Adler, A. F., Kabadi, A. M., Polstein, L. R., et al. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* 10, 973–976. doi: 10.1038/nmeth.2600

Rada, C., and Di Noia, J. M. (2004). Mismatch recognition and uracil excision provide complementary paths to both Ig switching and the A/T-focused phase of somatic mutation. *Mol. Cell* 16, 163–171. doi: 10.1016/j.molcel.2004.10.011

Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., and Trevino, A. E. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154, 1380–1389. doi: 10.1016/j.cell.2013.08.021

Sachse, C., Krausz, E., Krönke, A., Hannus, M., Walsh, A., and Grabner, A. (2005). High-throughput RNA interference strategies for target discovery and validation by using synthetic short interfering RNAs: functional genomics investigations of biological pathways. *Methods Enzymol.* 392, 242–277. doi: 10.1016/S0076-6879(04)92015-0

Salts, Y., Beckmann, J. S., Loyter, A., and Lavi, U. (1985). Interactions of sendai virus with plant protoplasts. *Plant Sci.* 41, 141–149. doi: 10.1016/0168-9452(85)90116-5

Sander, J. D., Cade, L., Khayter, C., Reyon, D., Peterson, R. T., and Joung, J. K. (2011). Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nat. Biotechnol.* 29, 697–698. doi: 10.1038/nbt.1934

Schiml, S., Fauser, F., and Puchta, H. (2014). The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. *Plant J.* 80, 1139–1150. doi: 10.1111/tpj.12704

Schunder, E., Rydzewski, K., Grunow, R., and Heuner, K. (2013). First indication for a functional CRISPR/Cas system in *Francisella tularensis*. *Int. J. Med. Microbiol.* 303, 51–60. doi: 10.1016/j.ijmm.2012.11.004

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., and Liang, Z. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 686–688. doi: 10.1038/nbt.2650

Shan, Q., Zhang, Y., Chen, K., Zhang, K., and Gao, C. (2015). Creation of fragrant rice by targeted knockout of the OsBADH2 gene using TALEN technology. *Plant Biotechnol. J.* 13, 791–800. doi: 10.1111/pbi.12312

Shmakov, S., Abudayyeh, O. O., Makarova, K. S., Wolf, Y. I., Gootenberg, J. S., and Semenova, E. (2015). Discovery and functional characterization of diverse class 2 CRISPR-cas systems. *Mol. Cell* 60, 385–397. doi: 10.1016/j.molcel.2015.10.008

Shukla, V. K., Doyon, Y., Miller, J. C., DeKelver, R. C., Moehle, E. A., and Worden, S. E. (2009). Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature* 459, 437–441. doi: 10.1038/nature07992

Sonoda, E., Hochegger, H., Saberi, A., Taniguchi, Y., and Takeda, S. (2006). Differential usage of non-homologous end-joining and homologous recombination in double strand break repair. *DNA Repair.* 5, 1021–1029. doi: 10.1016/j.dnarep.2006.05.022

Soyk, S., Müller, N. A., Park, S. J., Schmalenbach, I., Jiang, K., Hayama, R., et al. (2016). Variation in the flowering gene *SELF PRUNING 5G* promotes day-neutrality and early yield in tomato. *Nat. Genet.* 49, 162–168. doi: 10.1038/ng.3733

Stadler, L. J. (1928a). Mutations in barley induced by X-rays and radium. *Science* 68, 186–187.

Stadler, L. J. (1928b). Genetic effects of x-rays in maize. *Proc. Natl. Acad. Sci. U.S.A.* 14, 69–75.

Stroud, D. A., Surgenor, E. E., Formosa, L. E., Reljic, B., Frazier, A. E., and Dibley, M. G. (2016). Accessory subunits are integral for assembly and function of human mitochondrial complex I. *Nature* 538, 123–126. doi: 10.1038/nature19754

Suzuki, T., Eiguchi, M., Kumamaru, T., Satoh, H., Matsusaka, H., and Moriguchi, K. (2008). MNU-induced mutant pools and high performance TILLING enable finding of any gene mutation in rice. *Mol. Genet. Genomics* 279, 213–223. doi: 10.1007/s00438-007-0293-2

Svitashev, S., Young, J. K., Schwartz, C., Gao, H., Falco, S. C., and Cigan, A. M. (2015). Targeted mutagenesis, precise gene editing, and site-specific gene insertion in maize using Cas9 and guide RNA. *Plant Physiol.* 169, 931–945. doi: 10.1104/pp.15.00793

Symington, L. S., and Gautier, J. (2011). Double-strand break end resection and repair pathway choice. *Annu. Rev. Genet.* 45, 247–271. doi: 10.1146/annurev-genet-110410-132435

Takahashi, J. S., Pinto, L. H., and Vitaterna, M. H. (1994). Forward and reverse genetic approaches to behavior in the mouse. *Science* 264, 1724–1733. doi: 10.1126/science.8209253

Talamè, V., Bovina, R., Sanguineti, M. C., Tuberosa, R., Lundqvist, U., and Salvi, S. (2008). TILLMore, a resource for the discovery of chemically induced mutants in barley. *Plant Biotechnol. J.* 6, 477–485. doi: 10.1111/j.1467-7652.2008.00341.x

Townsend, J. A., Wright, D. A., Winfrey, R. J., Fu, F., Maeder, M. L., and Joung, J. K. (2009). High-frequency modification of plant genes using engineered zinc-finger nucleases. *Nature* 459, 442–445. doi: 10.1038/nature07845

Travella, S., Klimm, T. E., and Keller, B. (2006). RNA interference-based gene silencing as an efficient tool for functional genomics in hexaploid bread wheat. *Plant Physiol.* 142, 6–20. doi: 10.1104/pp.106.084517

Tsai, S. Q., Wyvekens, N., Khayter, C., Foden, J. A., Thapar, V., and Reyon, D. (2014). Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* 32, 569–576. doi: 10.1038/nbt.2908

Tsutsui, H., and Higashiyama, T. (2017). pKAMA-ITACHI vectors for highly efficient CRISPR/Cas9-mediated gene knockout in *Arabidopsis thaliana*. *Plant Cell Physiol.* 58, 46–56. doi: 10.1093/pcp/pcw191

Upadhyay, S. K., Kumar, J., Alok, A., and Tuli, R. (2013). RNA-guided genome editing for target gene mutations in wheat. *G3* 3, 2233–2238. doi: 10.1534/g3.113.008847

Vestergaard, G., Garrett, R. A., and Shah, S. A. (2014). CRISPR adaptive immune systems of Archaea. *RNA Biol.* 11, 156–167. doi: 10.4161/rna.27990

Wang, N., Wang, Y., Tian, F., King, G. J., Zhang, C., and Long, Y. (2008). A functional genomics resource for *Brassica napus*: development of an EMS mutagenized population and discovery of FAE1 point mutations by TILLING. *New Phytol.* 180, 751–765. doi: 10.1111/j.1469-8137.2008.02619.x

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., and Gao, C. (2014). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol.* 32, 947–951. doi: 10.1038/nbt.2969

Wei, C., Liu, J., Yu, Z., Zhang, B., Gao, G., and Jiao, R. (2013). TALEN or Cas9 - rapid, efficient and specific choices for genome modifications. *J. Genet. Genomics* 40, 281–289. doi: 10.1016/j.jgg.2013.03.013

Weinthal, D. M., Taylor, R. A., and Tzfira, T. (2013). Nonhomologous end joining-mediated gene replacement in plant cells. *Plant Physiol.* 162, 390–400. doi: 10.1104/pp.112.212910

Wiedenheft, B., Sternberg, S. H., and Doudna, J. A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482, 331–338. doi: 10.1038/nature10886

Wright, A. V., Nuñez, J. K., and Doudna, J. A. (2016). Biology and applications of CRISPR Systems: Harnessing nature's toolbox for genome engineering. *Cell* 164, 29–44. doi: 10.1016/j.cell.2015.12.035

Wu, J. L., Wu, C., Lei, C., Baraoidan, M., Bordeos, A., and Madamba, M. R. (2005). Chemical- and irradiation-induced mutants of indica rice IR64 for forward and reverse genetics. *Plant Mol. Biol.* 59, 85–97. doi: 10.1007/s11103-004-5112-0

Wu, X., Scott, D. A., Kriz, A. J., Chiu, A. C., Hsu, P. D., and Dadon, D. B. (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat. Biotechnol.* 32, 670–676. doi: 10.1038/nbt.2889

Xiao, A., Wang, Z., Hu, Y., Wu, Y., Luo, Z., and Yang, Z. (2013). Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Res.* 41:e141. doi: 10.1093/nar/gkt464

Xie, K., Zhang, J., and Yang, Y. (2014). Genome-wide prediction of highly specific guide RNA spacers for CRISPR-Cas9-mediated genome editing in model plants and major crops. *Mol. Plant* 7, 923–926. doi: 10.1093/mp/ssu009

Xing, H. L., Dong, L., Wang, Z. P., Zhang, H. Y., Han, C. Y., and Liu, B. (2014). A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol.* 14:327. doi: 10.1186/s12870-014-0327-y

Xu, R., Qin, R., Li, H., Li, D., Li, L., Wei, P., et al. (2016). Generation of targeted mutant rice using a CRISPR-Cpf1 system. *Plant Biotechnol. J.* doi: 10.1111/pbi.12669. [Epub ahead of print].

Yang, H., Gao, P., Rajashankar, K. R., and Patel, D. J. (2016). PAM-dependent target DNA recognition and cleavage by C2c1 CRISPR-Cas endonuclease. *Cell* 167, 1814–1828. doi: 10.1016/j.cell.2016.11.053

Yang, L., Briggs, A. W., Chew, W. L., Mali, P., Guell, M., Aach, J., et al. (2016). Engineering and optimising deaminase fusions for genome editing. *Nat. Commun.* 7:13330. doi: 10.1038/ncomms13330

Zamore, P. D., Tuschl, T., Sharp, P. A., and Bartel, D. P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* 101, 25–33. doi: 10.1016/S0092-8674(00)80620-0

Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., and Essletzbichler, P. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163, 759–771. doi: 10.1016/j.cell.2015.09.038

Zhang, F., Maeder, M. L., Unger-Wallace, E., Hoshaw, J. P., Reyon, D., and Christian, M. (2010). High frequency targeted mutagenesis in *Arabidopsis thaliana* using zinc finger nucleases. *Proc. Natl. Acad. Sci. U.S.A.* 107, 12028–12033. doi: 10.1073/pnas.0914991107

Zhang, Y., Zhang, F., Li, X., Baller, J. A., Qi, Y., and Starker, C. G. (2013). Transcription activator-like effector nucleases enable efficient plant genome engineering. *Plant Physiol.* 161, 20–27. doi: 10.1104/pp.112.205179

Zhao, Y., Zhang, C., Liu, W., Gao, W., Liu, C., and Song, G. (2016). An alternative strategy for targeted gene replacement in plants using a dual-sgRNA/Cas9 design. *Sci Rep* 6:23890. doi: 10.1038/srep23890