Genome Editing—Principles and Applications for Functional Genomics Research and Crop Improvement

Hui Zhang, Jinshan Zhang, Zhaobo Lang, José Ramón Botella & Jian-Kang Zhu


To link to this article: https://doi.org/10.1080/07352689.2017.1402989

© 2017 The Author(s). Published with license by Taylor & Francis © Hui Zhang, Jinshan Zhang, Zhaobo Lang, José Ramón Botella and Jian-Kang Zhu

Published online: 11 Dec 2017.

Article views: 1310

View related articles

View Crossmark data

Citing articles: 1 View citing articles
I. Introduction

With the growth of the world’s population, crop production will need to be doubled by 2050 (Tilman et al., 2011). However, the current annual growth rate of the main crops, wheat, rice, soybean, and maize are 0.9%, 1.0%, 1.3%, and 1.6%, respectively, falling far behind the required 2.4% rate (Ray et al., 2013). Other factors, such as reduced arable land and water availability, climate change and increased demand for bio-fuels will further compound the problem in the future. Besides crop yield, there is an increasing demand to develop new varieties with improved traits, such as increased quality, enhanced nutrition, disease resistance, stress tolerance and reduced resource requirements. Traditional breeding approaches for crop improvement rely heavily on genetic variation that occurs spontaneously in nature or is artificially generated by chemical and physical mutagenesis using chemical mutagens and irradiation, or more modern approaches such as insertional mutagenesis by T-DNA insertions, or transposon tagging (Songstad et al., 2017). However, natural genetic variations may be limited, and the artificial mutagenesis methods have obvious disadvantages, such as the random nature of the induced mutations, their low efficiency as well as being time-consuming, laborious and costly. The development of gene targeting technology based on homologous recombination allowed the production of precise mutations, but the frequency of targeted integration was initially very low and only worked in a limited number of species, such as tobacco and rice (Puchta and Fauser, 2013). In recent years, emerging genome editing technologies have shown potential to revolutionize crop improvement making it possible to create new varieties in a fast, efficient and technically simple way. Most importantly, the ‘edited’ varieties can be free of transgenes and indistinguishable from those obtained using traditional breeding technologies (Songstad et al., 2017). Genome editing is also an extremely valuable tool for functional genomics research.

The term “genome editing” refers to technologies which can produce genome modifications, such as targeted mutagenesis or site-directed insertion/deletion/substitution, at specific sites in the genome of living organisms. Genome editing relies on the production of site-specific double-strand DNA breaks (DSBs) and the subsequent endogenous repair through the error-prone non-homologous end-joining (NHEJ) or the error-free homology-directed repair
(HDR) pathways. In plants, DSBs are mainly repaired by NHEJ, in which a variety of enzymes are used to directly join the break ends of the DSBs without the need for a homologous repair template (Puchta, 2005). NHEJ occurs throughout the cell cycle in higher eukaryotes and exhibits low fidelity in the repair (Mladenov and Iliakis, 2011). Owing to its error-prone nature, NHEJ repair often leads to the addition or deletion of nucleotides and, thus, DNA sequence alterations at the targeted DSB sites. In many cases, NHEJ can lead to a complete loss of gene function, as indels introduced in exons can lead to missense or nonsense mutations. So far, most of the published plant genome editing work has used the NHEJ pathway to knock-out genes. In the HDR pathway, a homologous sequence serves as a template to repair the DSBs allowing an accurate repair (Puchta, 2005). HDR can be used to introduce precise nucleotide sequence modifications or gene replacement/insertion at target loci in the presence of an exogenously supplied donor DNA as a repair template. Unlike NHEJ, HDR occurs mainly during the S/G2 phases of the cell cycle and has a much lower efficiency, making it more challenging in plants (Puchta, 2005). By taking advantage of the intrinsic DNA repair machinery of cellular organisms, tools that produce DSBs can be used to precisely alter the genome. So far, zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9 (CRISPR/Cas9) and CRISPR/CRISPR from Prevotella and Francisella 1 (CRISPR/Cpf1) system are the four primary genome editing tools used to produce site specific DSBs (Figure 1). In general, these approaches use sequence-specific nucleases (SSNs) composed of a DNA-binding domain to provide sequence specificity linked to a nuclease domain to introduce DNA strand breaks at the targeted sequence.

In this review, we will introduce the four available types of genome editing tools for inducing targeted DSBs, evaluate the advantages and disadvantages of each method and examine the practical applications of these tools for functional genomics research and crop improvement.

II. Principles of genome editing technologies

A. ZFN

ZFNs were the first generation of genome editing tools being first exploited to edit plant genomes in 2005 (Lloyd

---

**Figure 1.** Theoretical and practical application of SSN-based genome editing. (a) SSN-mediated site-specific double strands break (DSB). (b) Possible genome editing outcomes using SSNs; (c) Examples of genome editing outcomes for crop improvement.
ZFNs are engineered nucleases generated by fusing an artificial, sequence-specific zinc finger DNA-binding domain to a nonspecific DNA cleavage domain derived from the type II restriction endonuclease FokI (Kim et al., 1996). The DNA-binding domain contains several linked zinc finger (ZF) motifs, each of which recognizes a 3-bp specific DNA sequence (Figure 2a). Individual ZFNs typically contain three to six ZFs and can recognize a 9–18 bp long specific DNA sequence. The FokI nuclease must dimerize to cleave double-stranded DNA. Therefore, it is necessary to design a pair of ZFNs to produce a DSB at the desired DNA location (Bitinaite et al., 1998). To achieve FokI dimerization, two individual ZFNs need to bind to the forward and reverse strands respectively and the two target sequences, forward and reverse, must be separated by a 5 to 7 bp spacer sequence.

Genome modification using ZFNs have been successfully in various organisms including plants, such as corn,Figure 2. (For figure legend, see page 294.)
tobacco, Arabidopsis, and soybean, although not all engineered ZFNs can create DSBs efficiently (Shukla et al., 2009; Townsend et al., 2009; Zhang et al., 2010; Sander et al., 2011). The design and assembly of ZFNs are technically challenging and outsourcing the construction of ZFN modules to commercial suppliers is expensive (Ramírez et al., 2008). These shortcomings have greatly limited its widespread adoption by the scientific community.

B. TALEN

TALENs belong to the second generation of genome editing tools and were first used for plant genome editing in 2011 (Cermak et al., 2011; Mahfouz et al., 2011). Like ZFNs, TALENs are engineered nucleases although in this case they are generated by fusing a transcription activator-like effector (TALE) DNA binding protein to the non-specific DNA endonuclease FokI (Miller et al., 2011). TAL effector proteins (TALEs) are secreted by Xanthomonas bacteria upon infection of their host plants and bind to the plant DNA through a domain containing several tandem 34–35 amino-acid repeats (Boch and Bonas, 2010). A typical TALE protein contains an N-terminal translocation signal, a C-terminal acidic transcription-activation domain and a central DNA-binding domain (Bogdanove et al., 2010) (Figure 2b). The DNA-binding domain is composed of several tandem of nearly identical 33–35 amino acid repeats which are highly conserved except for two adjacent residues (positions 12 and 13) named repeat variable di-residues or RVDs (Bogdanove et al., 2009). In 2009, two independent groups simultaneously broke the code of TALE’s DNA binding specificity (Boch et al., 2009; Moscou and Bogdanove, 2009). Each of the individual repeat units in the DNA-binding domain binds to a single nucleotide, and the RVDs in the repeats show a strong correlation with a specific nucleotide (A, G, C, or T). For example, the RVDs His-Asp (HD), Asn-Ile (NI), Asn-Gly (NG) and Asn-Asn (NN) recognize the nucleotides C, A, T, and G, respectively (Boch et al., 2009; Moscou and Bogdanove, 2009; Deng et al., 2012a). This straightforward relationship between protein sequence and DNA recognition allows the engineering of specific DNA-binding proteins by selecting specific combinations of the four prominent repeats containing the appropriate RVDs. Thus, TALENs can be designed and assembled to target any DNA sequence, although they are sensitive to methylation (Deng et al., 2012b; Kaya et al., 2017b). As was the case for ZFNs, two TALEN monomers are required to form a dimerized functional FokI. The assembly for TALENs are easier than ZFNs, resulting in a wider adoption of the editing method in plants, including Arabidopsis, rice, Brachypodium, barley, maize, tobacco, soybean, wheat, tomato, potato and sugarcane (Cermak et al., 2011; Mahfouz et al., 2011; Li et al., 2012a; Li et al., 2012b; Christian et al., 2013; Shan et al., 2013b; Wendt et al., 2013; Zhang et al., 2013; Haun et al., 2014; Liang et al., 2014; Lor et al., 2014; Wang et al., 2014; Nicolia et al., 2015; Jung and Altpeter, 2016).

C. CRISPR/Cas9

CRISPR/Cas9 is regarded as the third-generation genome-editing tool. It was first used to edit plant genes in 2013 and is currently the prevalent gene editing tool (Li et al., 2013; Nekrasov et al., 2013; Shan et al., 2013a). CRISPR was discovered as a prokaryotic immune system that protects cells by selectively targeting and destroying foreign DNA, such as viruses or plasmids (Horvath and Barrangou, 2010; Marraffini and Sontheimer, 2010). The engineered CRISPR/Cas9 is based on the type II CRISPR system that has the following three main components: the CRISPR associated protein 9 (Cas9), and two non-coding CRISPR RNAs (crRNAs): a trans-activating

Figure 2. (see previous page) Schematic models of the gene editing systems. (a) Schematic of the ZFN system. Each zinc finger specifically recognizes a 3-bp sequence. Paired ZFNs bind to the opposite strands to dimerize FokI, producing a DSB at desired site. (b) Schematic of the TALEN system. Engineered TALE proteins contain an N-terminal region, a central domain of repeats and a C-terminal region. The central repeat domain is composed of several nearly identical tandem repeats of 33–35 amino acids. The two adjacent residues in positions 12 and 13 in each repeat, designated RVDs, break the conservation of the repeats and provide specificity for a specific nucleotide A, G, C, or T. Paired TALENs bind to the opposite strands to dimerize FokI, which produces a DSB at desired site. (c) Schematic of the CRISPR/Cas9 system. The engineered CRISPR/Cas9 system consists of two components: a single-guide RNA (sgRNA) and the Cas9 endonuclease. The sgRNA contains a spacer sequence followed by 79-nt of an artificially fused tracrRNA and crRNA sequence. The spacer sequence is typically 20-nt and specifically binds to the target DNA sequence containing a 5’-NGG-3’ PAM motif at the 3’ end. The fused tracrRNA and crRNA sequence forms a stem-loop RNA structure that binds to the Cas9 enzyme. Cas9 uses its HNH nuclease domain to cleave the DNA strand complementary to the spacer sequence and its RuvC-like nuclease domain to cleave the DNA strand non-complementary to the target, creating a blunt end DSB about 3–4 nucleotides upstream of the PAM sequence. (d) Schematic of the CRISPR/Cpf1 system. The engineered CRISPR/Cpf1 system consists of two components: crRNA and Cpf1. The crRNA contains a spacer sequence, typically 23-nt followed by a 19-nt repeat sequence. The spacer binds to the DNA sequence, which needs to have a 5’-TTN-3’ PAM motif at its 5’ region, while the repeat sequence forms a stem-loop RNA structure that binds to the Cpf1 enzyme. Cpf1 uses its RuvC-like nuclease domain to cleave the DNA strand non-complementary to the target 17–18 nt downstream of the PAM sequence and a putative novel nuclease domain Nuc to cleave the DNA strand complementary to the target at 22–23 nt downstream of the PAM sequence creating a DSB with sticky ends.
crRNA (tracrRNA), and a precursor crRNA (pre-crRNA) (Horvath and Barrangou, 2010; Bhaya et al., 2011). Cas9 is a DNA endonuclease which contains an HNH nuclease domain and a RuvC-like nuclease domain and is involved in the crRNA maturation process and crRNA-guided DNA cleavage (Horvath and Barrangou, 2010; Bhaya et al., 2011). The tracrRNA is a small trans-encoded RNA containing a sequence with almost perfect complementarity to the repeats within the pre-crRNA to allow the formation of an RNA duplex necessary for crRNA maturation and crRNA-guided DNA cleavage (Horvath and Barrangou, 2010; Bhaya et al., 2011). The pre-crRNAs are transcribed from CRISPR loci, which consist of a recurring repeat-spacer array. The repeats (usually between 23 and 47 bp) are typically identical in length and sequence within a CRISPR locus, but varies greatly among different loci. Most repeat sequences show palindromes or are short inverted repeats that can form hairpin-shaped secondary structures (Horvath and Barrangou, 2010). The spacers (typically 21–72 bp) are derived from invading viral or plasmid DNA and can guide Cas9 to cleave an invading protospacer, the sequence in the foreign genome from which spacers are derived, on subsequent invasion by viruses or plasmids (Horvath and Barrangou, 2010; Bhaya et al., 2011). The spacer sequences are typically unique within a CRISPR locus and the size is similar to that of the repeats in the same array (Grissa et al., 2007). The pre-crRNA encompasses much of the CRISPR repeat-spacer array and is transcribed together with the tracrRNA. Subsequently, the tracrRNA hybridizes with the pre-crRNA to form an RNA duplex and associates with Cas9. The duplex is then targeted by RNase III to produce the mature crRNAs with a truncated spacer at one end. The mature crRNA:tracrRNA duplex, mainly the 20 nucleotides at the 5’ end of the crRNA, directs Cas9 to the DNA target sequence consisting of protospacer adjacent motifs (PAM) and complementary protospacer sequence. Finally, the Cas9 HNH nuclease domain cleaves the DNA strand that is complementary to the RNA guide while the RuvC-like nuclease domain cleaves the DNA strand that is non-complementary to the target to create a DSB within the protospacer about 3–4 nucleotides upstream of the PAM (Horvath and Barrangou, 2010; Bhaya et al., 2011; Cong et al., 2013).

Engineered CRISPR/Cas9 systems have been developed based on the type II CRISPR to induce sequence-specific DSBs and targeted genome editing (Jinek et al., 2012; Cong et al., 2013; Mali et al., 2013). In this system, the original crRNA and tracrRNA have been fused into one single guide RNA (sgRNA) for technical simplicity. In this way only two components are necessary for CRISPR/Cas9 function: the Cas9 endonuclease, which is used to cleave the target sequence, and a sgRNA, which defines the specificity of the Cas9 and guides Cas9 to the target DNA (Cong et al., 2013; Mali et al., 2013). In the engineered system, programming Cas9 to target a specific genome sequence only requires the introduction of a 20-bp spacer sequence in the sgRNA (Figure 2c). Therefore, CRISPR/Cas9 is much simpler and easier to manipulate than the previously available methods ZFNs and TALENs, and most importantly, it has a much higher efficiency in producing targeted mutations. Multiple sgRNAs with different target sequences can be easily designed for simultaneous multiplex editing. CRISPR’s technical simplicity and efficiency made it a true breakthrough in genome editing, allowing non-specialized laboratories to use it routinely. Since its first report for plant gene editing in 2013, CRISPR/Cas9 has been used in a large number of species, including rice, wheat, tobacco, Arabidopsis, sorghum, tomato, maize, potato, poplar, soybean, barley, moss, Brassica oleracea, sweet orange, apple, liverwort, grape, lettuce, cotton, Lotus japonicus, dandelion, flax, petunia, citrus, watermelon and mushroom (Jiang et al., 2013; Li et al., 2013; Mao et al., 2013; Nekrasov et al., 2013; Shan et al., 2013a; Brooks et al., 2014; Feng et al., 2014; Jia and Wang, 2014; Sugano et al., 2014; Fan et al., 2015; Lawrenson et al., 2015; Li et al., 2015; Svitashev et al., 2015; Wang et al., 2015b; Woo et al., 2015; Iaffaldano et al., 2016; Lopez-Obando et al., 2016; Nishitani et al., 2016; Ren et al., 2016; Sauer et al., 2016; Waltz, 2016; Wang et al., 2016b; Zhang et al., 2016a; Chen et al., 2017; Jia et al., 2017; Tian et al., 2017).

D. CRISPR/Cpf1

CRISPR/Cpf1 is also regarded as a third-generation genome editing tool, and it adds another option to the CRISPR toolbox. The CRISPR/Cpf1 system was adopted as a genome editing tool in 2015 and used for plant gene editing in 2016 (Zetsche et al., 2015; Endo et al., 2016a). CRISPR/Cpf1 contain two main components, the Cpf1 enzyme and the crRNA which determines the specificity of the system (Figure 2d). Although the CRISPR/Cpf1 and CRISPR/Cas9 systems are similar, but also have some important differences (Shmakov et al., 2017). First, the CRISPR/Cpf1 system does not need a trans-acting crRNA (tracrRNA), which is necessary for crRNA maturity in the CRISPR/Cas9 system. The Cpf1-crRNA complex can efficiently cleave the target double-strand DNA. Second, the CRISPR/Cpf1 engineered crRNA is about 42–44 nt long, including a 19 nt repeat and a 23–25 nt spacer, compared to the ~100 nt sgRNA in the CRISPR/Cas9 system (Zetsche et al., 2015) (Figures 2c and 2d). Third, in addition to the nuclease activity to cleave
double stranded DNA, Cpf1 also functions as an RNase to process the pre-crRNA arrays into mature single crRNAs (Zetsche et al., 2015; Dong et al., 2016; Fonfara et al., 2016). This makes the CRISPR/Cpf1 system, which requires only one promoter to drive an array of several small crRNAs, an even simpler tool than CRISPR/Cas9, which needs to separately express individual sgRNAs when attempting to edit multiple targets/genes. Thus requiring the use of multigene cassettes. Fourth, unlike Cas9 containing RuvC and HNH nuclease domains, which cuts both DNA strands in the same position (3–4 bp upstream of the PAM) to produce blunt ends, Cpf1 contains one RuvC-like domain and a novel nuclease domain that cleave the target sequence 23 bp and the non-target strand 18 bp downstream of the PAM sequence, producing a sticky end with 5-bp overhangs (Zetsche et al., 2015) (Figure 2c). This property makes Cpf1-induced mutations usually larger than Cas9-induced mutations, which typically are 1-bp indel mutations. The resulting sticky ends should, in theory, increase the efficiency of HDR-mediated donor DNA insertion in the Cpf1 cleaved site (Zetsche et al., 2015).

Fifth, while CRISPR/Cas9 requires a G-rich (5′-NGG-3′) PAM sequence at the 5′ end of the target sequence, CRISPR/Cpf1 requires a T-rich (5′-TTTN-3′ or 5′-TTN-3′) PAM sequence located at the 5′ end of the target sequence (Zetsche et al., 2015). Currently, three engineered CRISPR/Cpf1 systems have been developed, including FnCpf1 from Francisella novicida, AsCpf1 from Acidaminococcus sp. and LbCpf1 from Lachnospiraceae bacterium. All three Cpf1 systems have been used for plant genome editing in several species, including rice, Arabidopsis, tobacco, and soybean (Endo et al., 2016a; Kim et al., 2017; Tang et al., 2017; Wang et al., 2017a; Xu et al., 2017).

III. Comparison of the four types of genome editing tools

A. Sequence selection, assembly, and cost

Engineering ZFNs rely on dimerized FokI to cleavage target DNA, and the target specificity is determined largely by ZFs (Kim et al., 1996). ZF modules have been developed to recognize all of the 64 (4^4*4) possible nucleotide triplets (Carroll et al., 2006). Theoretically, ZFNs can be designed to target any sequence. However, the specificities of individual zinc fingers depend on the context of the surrounding zinc fingers and DNA, and no method can account for this context dependence (Ramirez et al., 2008). Therefore, in practice, there is no guarantee that a suitable ZFN target sequence can be found for a specific gene or chromosomal loci. The assembly of ZFNs is complicated, involving laborious and time-consuming steps (Carroll et al., 2006; Gonzalez et al., 2010), and the assembled ZFNs frequently fail to cleave target sites (Ramirez et al., 2008). Finally, the commercial assembly of ZFNs was prohibitively expensive, which has greatly restricted their use in most laboratories.

TALENs also need dimerized FokI to cleavage target DNA, and TALE repeat domains determine the sequence specificity of cleavage (Miller et al., 2011). TALEN target selection is limited by the requirement that TALE binding sites should start with a thymidine residue (T) and it is sensitive to DNA methylation (Cermak et al., 2011; Valton et al., 2012). Owing to the simple one-to-one specific recognition relationship between TALE repeats and the four nucleotides, A, T, C and G, TALENs are easier to design and assemble than ZFNs, but the assembly is still quite laborious and time-consuming (Cermak et al., 2011). Not all TALENs are efficient to cleave target sites, and TALEN pairs must be experimentally validated. For example, in our previous work, only six of 10 TALEN constructs showed activity in transgenic T0 plants (Zhang et al., 2015a). Like ZFNs, custom-designed TALE arrays are commercially available but are also expensive. Both ZFNs and TALENs rely on engineering proteins to specifically recognize target sequences and the complicated assembly process is the major hurdle preventing their wider application for genome editing.

In comparison, CRISPR/Cas9 and CRISPR/Cpf1 rely on a single RNA molecule, sgRNA or crRNA, to direct Cas9 or Cpf1 to its complementary DNA sequence. For target selection, the CRISPR/Cas9 system requires the presence of an NGG PAM sequence downstream of the target sequence and CRISPR/Cpf1 requires a TTN or TTTN PAM sequence upstream of the target sequence. Despite the PAM limitation, it is easy to find target sequences using these two systems. For example, analysis of eight plant species showed that 5–12 NGG–PAMs can be found in every 100 bp of DNA sequence (Xie et al., 2014), and 3.5 times more TTN-PAM sites than NGG-PAM sites can be identified in Arabidopsis (Minkenberg et al., 2017). A large number of web-based tools have been developed to facilitate the design of specific gRNA spacers, such as CRISPR-PLANT (http://www.genome.arizona.edu/crispr/) and CRISPR-P (http://crispr.hzau.edu.cn/CRISPR/) (Lei et al., 2014). Cloning the necessary 20 bp or 23–25 bp gRNA spacers in Cas9 or Cpf1 cassettes respectively to target different sequences is technically easy. The 20 bp or 23–25 bp specific sequences can be synthetized as complementary oligonucleotides with specific adapters and then annealed before ligating them into the functional CRISPR/Cas9 or CRISPR/Cpf1 vector. This simplicity is the main reason for the widespread
adoption of the CRISPR/Cas9 and CRISPR/Cpf1 systems and their application in numerous organisms (Table 1).

### B. Multiplex editing

The ease of multiplexing is another main advantage of CRISPR/Cas9 and CRISPR/Cpf1 compared to ZFNs and TALENs (Table 1). Multiplex editing using ZFNs or TALENs requires the production of separate dimeric proteins to target each of the different loci. Taking into account the technical difficulty of assembling a single ZFN or TALEN, multiplexing is only theoretically possible. In contrast, CRISPR/Cas9 can efficiently target multiple sites by providing a Cas9 cDNA and multiple sgRNAs in a single expression cassette (Wang et al., 2013; Xie et al., 2015). Several approaches have been developed to express multiple sgRNAs, such as Golden Gate assembly, polycistronic tRNA–gRNA system, self-cleaving ribozyme flanked gRNAs and target-adaptor ligation (Gao and Zhao, 2014; Lowder et al., 2015; Ma et al., 2015; Xie et al., 2015). Compared to the multiple sgRNA cassettes required to multiplex CRISPR/Cas9, CRISPR/Cpf1 is even simpler and only needs a single, short repeat-spacer array for multiplex genome editing because Cpf1 not only cleaves target DNA but also processes repeat-spacer array to functional repeat-spacer (crRNA) units. For example, our group simultaneously mutated four genes with high efficiency using the CRISPR/Cpf1 system (Wang et al., 2017a). To construct this CRISPR/Cpf1 multiplexing editor, only two 176 bp complementary oligonucleotides containing four repeat-spacer repeats (4’42 bp), were synthesized with specific adapters (2’4 bp), annealed and ligated into the functional vector (Wang et al., 2017a).

### C. Efficiency

It is difficult to compare the efficiency of individual sequence-specific nucleases as in many occasions it depends on the selected target. In general, CRISPR/Cas9 has higher efficiency than ZFNs and TALENs, whereas TALENs usually exhibit higher efficiency than ZFNs (Table 1). From our own experience, nearly all designed CRISPR/Cas9 cassettes can generate cleavage in target sequences with high efficiency, but not all designed TALENs can produce the predicted mutations. Even after optimization by truncating the TALE portion of TALEN to improve efficiency, the efficiency was still lower than CRISPR/Cas9 (Zhang et al., 2014; Zhang et al., 2015a). From the limited published literature, it seems that CRISPR/Cpf1 efficiency is comparable to CRISPR/Cas9 (Wang et al., 2017a; Xu et al., 2017).

### D. Off-target effects

TALENs and ZFNs can produce some off-target effects, and in order to reduce off-targets, several approaches, such as using FokI variants and nickases, have been developed (Miller et al., 2007; Gabriel et al., 2011; Mussolino et al., 2011). Theoretically, increasing the number of ZFP and TALEs should improve targeting specificity by increasing the length of the target sequence. However, in practice, extended ZPF or TALE modules also increase the likelihood to bind off-target sites (Carroll, 2011).

CRISPR/Cas9 specificity relies on 22 bp target sequences composed of a 20-bp spacer sequence within sgRNA and a 5’-NGG-3’ PAM sequence recognized by Cas9 (Jinek et al., 2012; Wang et al., 2015a). The binding of sgRNA to the target sequence can tolerate mismatches of several nucleotides, thus increasing the possibility of off-target effects (Fu et al., 2013; Hsu et al., 2013; Pattanayak et al., 2013). Various strategies have reduced off-target effects, including selecting highly specific target sequences, optimizing nuclease expression, using truncated sgRNAs, dCas9-FokI fusions, and paired Cas9-nickases (Fuji et al., 2013; Hsu et al., 2013; Pattanayak et al., 2013; Cho et al., 2014; Fu et al., 2014; Guilinger et al., 2014; Tsai et al., 2014). Current evidence indicates that off-target effects are limited in CRISPR–Cas9 edited plants (Zhang et al., 2014; Zhang et al., 2016b). Studies in mammalian cells have shown that compared with CRISPR/Cas9, CRISPR-Cpf1 exhibited minimal off-target effects and from the limited published literature, it seems that is also the case in plants (Hur et al., 2016; Kim et al., 2016; Kleinstiver et al., 2016). The use of web-based tools can help with the selection of highly specific target sequences to minimize or even completely avoid off-target effects in genome editing (Lei et al., 2014; Montague et al., 2014; Lee et al., 2016; Rastogi et al., 2016). Nevertheless, although there are rising concerns about the off-target effects of genome editing applications in human therapy, it is not a big issue for plant research or breeding purposes because any unwanted

---

**Table 1. Comparison of the four types of genome editing tools.**

<table>
<thead>
<tr>
<th></th>
<th>ZFN</th>
<th>TALEN</th>
<th>CRISPR/Cas9</th>
<th>CRISPR/Cpf1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection of target sequence</td>
<td>Limited</td>
<td>Abundant</td>
<td>Very abundant</td>
<td>Very abundant</td>
</tr>
<tr>
<td>Assembly</td>
<td>Laborious</td>
<td>Very expensive</td>
<td>Simple</td>
<td>Simple</td>
</tr>
<tr>
<td>Cost</td>
<td>Expensive</td>
<td>Laborious</td>
<td>Cheap</td>
<td>Cheap</td>
</tr>
<tr>
<td>Multiplex editing</td>
<td>Difficult</td>
<td>Difficult</td>
<td>Easy</td>
<td>Very easy</td>
</tr>
<tr>
<td>Efficiency</td>
<td>Low</td>
<td>Moderate</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Off-target effects</td>
<td>Low</td>
<td>Low</td>
<td>Very good</td>
<td>Low</td>
</tr>
<tr>
<td>Overall evaluation</td>
<td>Good</td>
<td>Very good</td>
<td>Excellent</td>
<td>Excellent</td>
</tr>
</tbody>
</table>

---

CRITICAL REVIEWS IN PLANT SCIENCES 297
changes can be easily monitored and segregated out in subsequent generations.

**Overall evaluation**

In principle, all genome editing applications achieved by CRISPR/Cas9 or CRISPR/Cpf1 can also be accomplished using either ZFNs or TALENs. However, mainly due to their complicated assembly process and the near impossibility to multiplex, ZFNs and TALENs have been mostly abandoned in favor of the CRISPR/Cas9 system (Table 1). Nevertheless, the emerging CRISPR/Cpf1 system is establishing itself as a promising alternative to CRISPR/Cas9.

**IV. Applications of genome editing technologies in functional genomics research**

**Gene knockout**

1. **Single gene knockout**

At present, the most widely used and important application of genome editing is to knockout target genes. In plants, NHEJ is the main pathway used to repair DSBs, and the process can introduce small deletions or insertions (indels), typically smaller than 100 bp (Puchta, 2005). Introduction of indels in a coding region mostly leads to frameshift mutations resulting in the loss of gene function. Most importantly, the mutations are stable and heritable in future generations. Due to its simplicity and high efficiency, CRISPR/Cas9 is now the dominant tool for knocking out genes.

2. **Multiplex gene knockout**

Genome editing has been used for simultaneous targeting of multiple genes in many plant species, including Arabidopsis, rice, maize, soybean and tobacco (Li et al., 2013; Xie et al., 2015; Zhang et al., 2015b; Char et al., 2016; Minkenberg et al., 2016; Chilcoat et al., 2017; Wang et al., 2017a). For example, our group used CRISPR/Cas9 for multiplex gene editing to create a PYL/ PYR sextuple mutant, using a single construct harboring expression cassettes for six different sgRNAs (Zhang et al., 2015b). A homozygous sextuple mutant was obtained in the T3 generation exhibiting the expected phenotypes (Zhang et al., 2015b). Multiplex gene editing is not only useful for functional genomics research, such as the study of redundant gene families and functionally related genes but is also important for crop improvement, allowing fast pyramiding of multiple traits. For example, CRISPR/Cas9 was used to simultaneously knockout three negative regulators of grain size in rice, GW2, GWS, and TGW6 and the new varieties exhibited 20%–30% increases in grain size and weight compared to the wild type (Xu et al., 2016).

3. **Large fragment deletions**

When two DSBs are introduced on the same chromosome at a certain distance, the two sites may connect through the NHEJ pathway resulting in the deletion of the intervening sequence. Relatively large deletions are useful for some purposes in research and crop improvement, such as the study of gene clusters and noncoding RNAs. TALEN and CRISPR have been used to produce large deletions in species such as rice, Arabidopsis, and tobacco (Christian et al., 2013; Shan et al., 2013b; Ordon et al., 2017). In rice, up to 245 kb has been removed from the genome with a high frequency using CRISPR/Cas9 (Zhou et al., 2014) and our group successfully deleted a large genomic fragment in Arabidopsis containing the CBF1, CBF2 and CBF3 genes (Zhao et al., 2016).

4. **Gene knockout in polyploid plants**

Many crops are polyploid. Owing to their complex genetic composition, gene knockout mutants are quite difficult to obtain using traditional genetic approaches, thus hampering studies on gene function. Genome editing tools are especially useful for polyploid crops lacking mutant resources. Until now, genome editing technologies have been successfully applied for gene knockout in triploids (citrus and apple), tetraploids (cotton, pasta wheat and potato), hexaploids (Camelina and bread wheat), and octaploids (sugarcane) (D’Halluin et al., 2013; Wang et al., 2014; Clasen et al., 2015; Nicolia et al., 2015; Wang et al., 2015b; Butler et al., 2016; Forsyth et al., 2016; Jung and Altpeter, 2016; Malnoy et al., 2016; Morineau et al., 2016; Nishitani et al., 2016; Zhang et al., 2016b; Chen et al., 2017; Jia et al., 2017; Jiang et al., 2017; Li et al., 2017a; Liang et al., 2017; Peng et al., 2017). Some of the work has shown a strong potential for gene functional studies as well as practical crop improvement. For example, in Triticum aestivum (bread wheat) all six alleles of the MLO gene were mutated using TALENs and CRISPR/Cas9, which is considered nearly impossible using traditional breeding methods, rendering the new variety resistant to the fungal pathogen causing powdery mildew (Wang et al., 2014). It is expected that the same approach will be used in the near future in many polyploid crops to study gene function and to produce new varieties.

5. **Gene targeting**

Gene targeting refers to the use of genetic engineering methods to produce a one-for-one substitution of a DNA fragment (gene replacement) or the insertion of a
new sequence in a specific genomic locus (gene knock-in). Gene targeting has many applications in functional genomics research, such as precise gene modifications and epitope tagging of endogenous proteins. In addition, many agriculturally important traits are conferred by the gene coding region or promoter region, making gene targeting also useful for crop improvement. Gene targeting has been the focus of research for a long time, mostly based on homologous recombination, but the low frequency of targeted integration limited its use to a very few species such as tobacco and rice (Paszkowski et al., 1988; Terada et al., 2002). It was predicted that the gene targeting efficiency could increase up to 100-fold if a DSB exists at the target locus as this was expected to engage the cellular HDR repair mechanism, making it an obvious potential application for the use of genome editing tools (Puchta et al., 1996). In the last several years, ZFN, TALEN and CRISPR/Cas9 have been successfully used for gene targeting in tobacco, maize, Arabidopsis, tomato, rice, barley, flax, moss soybean and wheat (Shukla et al., 2009; Townsend et al., 2009; Zhang et al., 2013; Fauser et al., 2014; Budhagatapalli et al., 2015; Cermak et al., 2015; Li et al., 2015; Svitashev et al., 2015; Endo et al., 2016b; Li et al., 2016a; Li et al., 2016c; Sauer et al., 2016; Sun et al., 2016b; Collonnier et al., 2017; Gil–Humanes et al., 2017). One major obstacle for HDR-mediated gene targeting is its low efficiency, mostly because the NHEJ pathway activity repairing DSBs is two orders of magnitude higher than the HDR pathway (Steinert et al., 2016). It is therefore theoretically possible to increase HDR-mediated gene targeting efficiency by suppression of the NHEJ pathway. Indeed, Arabidopsis NHEJ pathway mutants ku70 and lig4 showed a marked improvement in ZFN-mediated knock-in efficiency (5-16-fold and 3-4-fold respectively) (Qi et al., 2013). Another way to improve HDR-mediated gene targeting is to deliver large amounts of repair template, donor DNA, to the plant nucleus. Particle bombardment can provide multiple copies of donor DNA and has been employed for genome editing-assisted gene targeting in multiple plants, including rice, cotton, soybean, barley and maize (D’Halluin et al., 2013; Budhagatapalli et al., 2015; Li et al., 2015; Svitashev et al., 2015; Endo et al., 2016b; Li et al., 2016c). Protoplasts can also be transfected with large amounts of donor DNA, and some of the genome editing-mediated knock-in work have been performed in protoplasts (Townsend et al., 2009; Zhang et al., 2013; Sauer et al., 2016). However, for most crops, regeneration of whole plants from protoplasts is quite challenging (Baltes et al., 2017). An alternative method to deliver abundant donor DNA is to use the geminivirus system, which has the property of excising a fragment of its genomic DNA once inside a cell to produce a self-replicating plasmid (Hanleybowdoin et al., 2013). Recently, geminiviral replicons were used to achieve high-efficiency gene targeting in tobacco, tomato, wheat, potato and rice (Baltes et al., 2014; Cermak et al., 2015; Butler et al., 2016; Gil–Humanes et al., 2017; Wang et al., 2017b). As an example, our group has achieved a remarkable 19.4% efficiency in targeted knock-in using a geminivirus-based CRISPR/Cas9 system in transgenic rice plants (Wang et al., 2017b). Gene targeting has also been reported in wheat through the NHEJ pathway using TALENs to produce DSBs (Wang et al., 2014). CRISPR/Cpf1 has been suggested to be more efficient than CRISPR/Cas9 for gene targeting via the NHEJ mechanism because of the staggered nature of the DNA strand breaks (Zetsche et al., 2015). Genome editing-mediated gene knock-in has already been used to produce many crop varieties with improved traits, such as herbicide resistance in rice, maize, and flax and increased drought resistance in maize, via introduction of precise point mutations, insertion of new genes or precise promoter replacement (Svitashev et al., 2015; Endo et al., 2016b; Li et al., 2016c; Sauer et al., 2016; Sun et al., 2016a; Shi et al., 2017).

6. Other applications

Precise base editing. Many important agronomic traits in crops are determined by variations in one or a few bases in the genomic sequence. The desired sequence can be introduced by HDR-mediated gene knock-in as discussed above. As an alternative, a CRISPR/Cas9-based Base Editor (BE) technology has been recently developed that enables the editing of a single base C to T, resulting in C→T (or G→A) substitution, without DSBs or exogenous donor DNA (Komor et al., 2016). This system uses a sgRNA and a modified Cas9. While the sgRNA is identical to the one used in the normal CRISPR/Cas9 system, the Cas9 has been modified and fused to a cytidine deaminase (CDA) which can convert cytidine to uridine. The modified Cas9 is either a deactivated Cas9 (dCas9, with D10A and D840A substitutions in its catalytic sites) or a nickase Cas9 (nCas9 with a D10A substitution). This system has been used in rice, wheat, maize and tomato to create C to T substitutions in a window from position 3 to 9 within the protospacer, counting from the end distal to the PAM as position 1 (Lu and Zhu, 2016; Li et al., 2017b; Shimatani et al., 2017; Zong et al., 2017). The APOBEC1 and PmCDAl cytidine deaminases from Rattus rattus (rat) and Petromyzon marinus (sea lamprey), respectively, have been successfully used in plants while the nCas9 has shown higher efficiency than the dCas9 in this system (Shimatani et al., 2017; Zong et al., 2017). Just like in human cells, fusion of
In the last several years, genome editing has been used to produce new crop varieties with improved traits, including increased yield, enhanced disease resistance, improved food quality and higher stress tolerance (Table 2).

### A. Improved yield

Grain yield is mainly determined by grain number, size and weight, all of which are typical quantitative traits, and many genes affecting crop yield have been characterized (Xing and Zhang, 2010; Bai et al., 2012). Knockout of genes known to negatively affect yield, such as GS3, DEP1, GS5, GW2, Gn1a, and TGW6 in rice, is a simple and direct way to improve crops. GS3, DEP1 and Gn1a have been individually mutated using CRISPR/Cas9, and some of the predicted phenotypes were observed (Li et al., 2016b; Shen et al., 2016). Simultaneous knockout of GW2, GW5, and TGW6 in rice resulted in a 29.8% increase in thousand-grain weight in the triple mutant (Xu et al., 2016). In bread wheat, thousand-kernel weight does not necessarily translate into improved crop yield, because large-scale field trials are necessary to verify the potential agronomic improvements.

<table>
<thead>
<tr>
<th>Crop</th>
<th>Editing tool</th>
<th>Target genes</th>
<th>Type of edit</th>
<th>Traits</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rice</td>
<td>CRISPR/Cas9</td>
<td>GW2, GW5 and TGW6</td>
<td>Gene knockout</td>
<td>Increased yield, increased thousand-kernel weight</td>
<td>Xu et al., 2016; Zhang et al., 2016b</td>
</tr>
<tr>
<td>Wheat</td>
<td>CRISPR/Cas9</td>
<td>GASR7</td>
<td>Gene knockout</td>
<td>Disease resistance</td>
<td>Li et al., 2012b</td>
</tr>
<tr>
<td>Citrus</td>
<td>TALEN</td>
<td>OsSWEET14</td>
<td>Promoter disruption</td>
<td>Disease resistance</td>
<td>Jia et al., 2017; Peng et al., 2017</td>
</tr>
<tr>
<td>Tobacco</td>
<td>CRISPR/Cas9</td>
<td>OsERF922</td>
<td>Gene knockout</td>
<td>Disease resistance</td>
<td>Wang et al., 2016a</td>
</tr>
<tr>
<td>Tobacco</td>
<td>TALEN</td>
<td>TaMLO</td>
<td>Gene knockout</td>
<td>Virus resistance</td>
<td>Wang et al., 2014</td>
</tr>
<tr>
<td>Cucumber</td>
<td>CRISPR/Cas9</td>
<td>eIF4E</td>
<td>Gene knockout</td>
<td>Virus resistance</td>
<td>Chandrasekaran et al., 2016</td>
</tr>
<tr>
<td>Tobacco</td>
<td>CRISPR/Cas9</td>
<td>Regions in the viral genome</td>
<td>Viral gene disruption</td>
<td>Virus resistance</td>
<td>Baltes et al., 2015; Ji et al., 2015</td>
</tr>
<tr>
<td>Tomato</td>
<td>CRISPR/Cas9</td>
<td>Regions in the viral genome</td>
<td>Viral gene disruption</td>
<td>Virus resistance</td>
<td>Ali et al., 2015</td>
</tr>
<tr>
<td>Tobacco</td>
<td>ZFN</td>
<td>ALS</td>
<td>HDR-mediated base change</td>
<td>Herbicide tolerance</td>
<td>Townsend et al., 2009</td>
</tr>
<tr>
<td>Maize</td>
<td>CRISPR/Cas9</td>
<td>ALS</td>
<td>HDR-mediated base change</td>
<td>Herbicide tolerance</td>
<td>Svitashhev et al., 2015</td>
</tr>
<tr>
<td>Soybean</td>
<td>CRISPR/Cas9</td>
<td>ALS</td>
<td>HDR-mediated base change</td>
<td>Herbicide tolerance</td>
<td>Li et al., 2016c</td>
</tr>
<tr>
<td>Rice</td>
<td>TALEN</td>
<td>ALS</td>
<td>HDR-mediated base change</td>
<td>Herbicide tolerance</td>
<td>Sauer et al., 2016</td>
</tr>
<tr>
<td>Rice</td>
<td>CRISPR/Cas9</td>
<td>EPSPS</td>
<td>HDR-mediated base change</td>
<td>Herbicide tolerance</td>
<td>Li et al., 2016a</td>
</tr>
<tr>
<td>Soybean</td>
<td>TALEN</td>
<td>FAD2-1A and FAD2-1B</td>
<td>Gene knockout</td>
<td>Improved oil composition</td>
<td>Haun et al., 2014</td>
</tr>
<tr>
<td>Soybean</td>
<td>TALEN</td>
<td>FAD2-1A and FAD2-1B and FAD3A</td>
<td>Gene knockout</td>
<td>Improved oil composition</td>
<td>Demorest et al., 2016</td>
</tr>
<tr>
<td>Camelina sativa</td>
<td>CRISPR/Cas9</td>
<td>FAD2</td>
<td>HDR-mediated base change</td>
<td>Herbicide tolerance</td>
<td>Morineau et al., 2016; Jiang et al., 2017</td>
</tr>
<tr>
<td>Potato</td>
<td>TALEN</td>
<td>VINV</td>
<td>Gene knockout</td>
<td>Reduced sugar content</td>
<td>Clasen et al., 2015</td>
</tr>
<tr>
<td>Rice</td>
<td>CRISPR/Cas9</td>
<td>TMS5</td>
<td>Gene knockout</td>
<td>Thermo-sensitive genic male sterile rice</td>
<td>Zhou et al., 2016</td>
</tr>
<tr>
<td>Maize</td>
<td>CRISPR/Cas9</td>
<td>Waxy</td>
<td>Gene knockout</td>
<td>Waxy corn</td>
<td>Chilcoat et al., 2017</td>
</tr>
<tr>
<td>Rice</td>
<td>CRISPR/Cas9</td>
<td>SC1 and SBEIb</td>
<td>Gene knockout</td>
<td>High amylose rice</td>
<td>Sun et al., 2017</td>
</tr>
</tbody>
</table>
B. Disease resistance

Plant diseases are the main cause of crop yield loss, and most importantly, diseases also affect produce quality for fresh consumption and food processing and safety (toxins) in many crops (Savary et al., 2012). Genome editing has been applied to increase disease resistance by editing disease-related genes. The rice OsSWEET14 is a host disease-susceptibility gene that is activated by a type-III effector protein secreted by the bacterial rice pathogen Xanthomonas oryzae pv. Oryzae causing bacterial blight (Antony et al., 2010). Disruption of the bacterial-derived protein binding sequence in the OsSWEET14 rice promoter using TALEN resulted in increased resistance to bacterial blight (Li et al., 2012b). CsLOB1 is also a host disease-susceptibility gene which plays a critical role in promoting pathogen growth and erumpent pustule formation in citrus (Hu et al., 2014). Recently, two groups generated canker-resistant citrus cultivars by CRISPR/Cas9-targeted modification of the CsLOB1 promoter (Jia et al., 2017; Peng et al., 2017). Knockout of the ERF transcription factor OsERF922, a negative regulator of rice blast resistance, resulted in enhanced resistance (Wang et al., 2016a). Editing of the wheat TaMLO gene is another good sample of the use of gene editing to introduce disease resistance into susceptible crop varieties (Wang et al., 2014). Loss-of-function mlo alleles in barley, Arabidopsis and tomato produce broad-spectrum and durable resistance to Blumeria graminis f. sp. tritici (Bgt) which cause powdery mildew (Wang et al., 2014). TALEN-induced mutation of all three TaMLO gene homoeologs produced heritable broad-spectrum resistance to powdery mildew in bread wheat (Wang et al., 2014). Using the same approach, CRISPR/Cas9-mediated gene disruption of the tomato SIMLO1 gene resulted in rapid generation of tomato fully resistant to powdery mildew (Nekrasov et al., 2017).

Aside from bacterial and fungal pathogens, plant viruses have also been targeted using genome editing approaches. The eukaryotic translation initiation factor eIF4E is a host factor required by plant RNA viruses to maintain their life cycle and mutations in this gene have produced broad virus resistance in T3 non-transgenic cucumbers (Chandrasekaran et al., 2016). Virus resistant tobacco and tomato have also been generated by directly targeting viral genomic sequences using CRISPR/Cas9 (Ali et al., 2015; Baltes et al., 2015; Ji et al., 2015).

C. Herbicide tolerance

Traditionally, herbicide tolerance in crops has been obtained through transgenesis (Lombardo et al., 2016). Genome editing provides a new approach to create herbicide tolerant crops and has been employed to edit endogenous plant genes, such as EPSPS and ALS, resulting in herbicide tolerant plants (Lombardo et al., 2016).

ALS encodes the acetolactate synthase enzyme that participates in the biosynthesis of branched-chain amino acids like valine, leucine, and isoleucine (Lee et al., 1988; Chipman et al., 1998). Inhibitors of ALS are used as herbicides that slowly starve affected plants of these amino acids, eventually leading to inhibition of DNA synthesis, but specific point mutations within the conserved region of ALS can confer resistance to these herbicides (Lee et al., 1988; Chipman et al., 1998; Svitashev et al., 2015). ALS is the target of numerous herbicides including sulfonylureas, imidazolinones, triazolopyrimidines, pyrrolizinyl oxybenzoates, and sulfonylamino carbonyl triazolinones (Zhou et al., 2007). Genome editing-based gene replacement has been used to introduce precise mutations in the ALS gene to produce herbicide tolerant plants, with the first example, tobacco, reported in 2009 using ZFNs and donor templates (Townsend et al., 2009). Herbicide-resistance maize, soybean, and rice have also been obtained using CRISPR/Cas9 and TALENs to introduce site-directed DNA base changes in the ALS gene (Li et al., 2015; Svitashev et al., 2015; Li et al., 2016c).

The EPSPS gene encodes a 5-enolpyruvylshikimate-3-phosphate synthase, which is necessary for the biosynthesis of aromatic amino acids essential for plant survival (Kishore and Shah, 1988). In plants, EPSPS is a target for glyphosate, a widely used herbicide which binds to EPSPS functional sites to prevent its activity (Kishore and Shah, 1988). The usual method to introduce glyphosate tolerance in plants is to modify the EPSPS protein structure in order to disrupt herbicide binding while maintaining its catalytic activity (Sammons and Gaines, 2014). CRISPR/Cas9 and single-stranded oligo DNA repair templates have been used in Linum usitatissimum (Flax) to substitute two nucleotides in the EPSPS glyphosate binding site through HDR-based genome editing (Sauer et al., 2016). The EPSPS edited flax showed higher levels of glyphosate tolerance than controls. A similar approach has been used to introduce base substitutions in the rice EPSPS gene resulting in glyphosate-resistant rice (Li et al., 2016a).

D. Healthy food

Genome editing can be used to modify plant components, resulting in healthier foods.

Improved oil composition

A high content of polyunsaturated fatty acids, particularly linolenic acid, in oils results in poor oxidative and
frying stability which limits their applications. Fatty acid desaturase (FAD) genes have been targeted to change fatty acid composition and improve oil quality. The FAD2 gene family is responsible for the conversion of oleic acid (monounsaturated) into linoleic acid while enzymes encoded by the FAD3 gene family catalyze the production of linolenic acid from linoleic acid (Demorest et al., 2016). TALENs were used to simultaneously knock out two soybean FAD2 genes, FAD2-1A and FAD2-1B, resulting in vastly improved oil quality: oleic acid increased from 20% to 80% and linoleic acid decreased from 50% to < 4% (Haun et al., 2014). To further improve oil composition, mutations in FAD3A were introduced into the previously produced fad2-1a/fad2-1b soybean plants by TALEN, resulting in further increased levels of oleic acid and decreased levels of linolenic acid (Demorest et al., 2016). Recently, two independent groups used CRISPR/Cas9 to simultaneously knock out all three FAD2 homeolog genes in the allohexaploid, *camelina sativa*, producing a significant enhancement in oil composition (Morineau et al., 2016; Jiang et al., 2017).

**Healthy potatoes**

Cold storage of potatoes reduces sprouting and ensures a continuous supply, but it also results in the accumulation of reducing sugars. The reducing sugars react with free amino acids during high temperature processing to produce brown, bitter-tasting products and increase the levels of acrylamide, which is a suspected human carcinogen and has caused global safety concerns (Clasen et al., 2015). VINV encodes a vacuolar invertase that catalyzes the conversion of sucrose into glucose and fructose, and has an essential role in the production of reducing sugars in cold-stored potato tubers. Mutation of VINV in a commercial Ranger Russet potato variety has been achieved using TALENs, with the resulting potatoes having undetectable levels of reducing sugars. Heat processing of the cold-stored potatoes resulted in reduced levels of acrylamide and produced lightly colored chips (Clasen et al., 2015).

**Other examples**

CRISPR/Cas9 targeted mutation of the *TMSS* gene in rice cultivars led to the rapid development of temperature-sensitive lines for use in hybrid rice production (Zhou et al., 2016). The maize *Waxy* (*Wx*) gene encodes a granule-bound starch synthase (GBSS) responsible for the synthesis of amylase in the kernel (Nelson and Rines, 1962). Wild type maize kernels consist of 75% amylopectin and 25% amylase while wx/wx lines contain nearly 100% amylopectin which is called waxy maize. The economically valuable waxy maize has been produced by CRISPR-mediated Waxy gene knockout (Chilcoat et al., 2017). High-amylose rice, with potential health advantages, was generated through CRISPR/Cas9-mediated knockout of the starch branching enzymes genes, *SBEI* and *SBEIIb* (Sun et al., 2017).

**VI. Concluding remarks and perspectives**

Overall, gene editing technologies, especially CRISPR/Cas9, have had a revolutionary influence on basic research in plants as well as crop improvement. One of the main advantages of these technologies is that the transgenes initially used to produce the genetic changes can be easily excised from the genome by genetic segregation, and the resulting gene-edited varieties are completely indistinguishable from those generated using conventional breeding methods. Recently, the US Department of Agriculture ruled that CRISPR-edited crops, including mushroom, and waxy corn were exempt from GMO regulation because they do not contain foreign DNA (Waltz, 2016). Application of DNA-free or integration-free genome editing approaches can further alleviate public concerns. Recently, preassembled Cas9 protein–gRNA ribonucleoproteins have been directly delivered into cells to successfully edit genes in several crops, such as rice and wheat (Woo et al., 2015; Svitashev et al., 2016; Liang et al., 2017). Preassembled Cpf1 protein–crRNA complexes have also been successfully used for plant genome editing (Kim et al., 2017). Using a different approach, the *Staphylococcus aureus* Cas9 (SaCas9) has been split into two halves that can spontaneously reassemble once expressed in plant cells (Kaya et al., 2017a). The reduced size of the fragments allows their cloning into viral vectors and thus a combination of a tomato mosaic virus-based vector and *Agrobacterium* has been used to transiently express both halves and the corresponding gRNA in *N. benthamiana* resulting in the production of targeted mutations (Kaya et al., 2017a). New and high-efficiency DNA-free genome editing approaches are expected to be developed in the near future and applied in food crops.

With the development of high-throughput sequencing technologies, many crop genomes such as walnut, apple, strawberry, grapevine, sweet orange, rice, maize, wheat, tomato, millet, etc., are becoming available, and the vast amount of genomic data accessible to researchers will facilitate gene functional studies as well as crop improvement through genome editing.

There are nevertheless important challenges for genome editing that needs to be overcome to facilitate their application in plants. Among those challenges is the development of highly efficient HDR-based genome
editing methods, which are especially useful for functional genomics and crop improvement. Efficient transformation and regeneration methods is another requisite that can limit gene editing applications in many crops. At present, many elite varieties as well as other important commercial crops are recalcitrant to transformation.

Although genome editing-based gene activation and repression have been applied in plants, other applications, such as epigenomic regulation also needs to be explored. Using genome editing to manipulate DNA methylation or histone modifications is also promising for basic research and crop improvement because the altered epigenetic marks may be inherited to future generations without changing the sequence of the genome itself. For example, in mammalian cells, CRISPR/dCas9 has been fused with DNA methyltransferase 3a (DNMT3a) and DNA demethylase (TET1) to induce DNA methylation or demethylation respectively in target regions (Liu et al., 2016). Although it is not yet available in plants, epigenome editing tools are expected to be developed in the near future because of their potential value.

Given the power of genome editing tools and the increasing number of researchers using and developing these tools, a revolutionary change is taking place in crop improvement that will help to meet the increasing demand for food and ensure world food security in the future.

Acknowledgments

Our work has been supported by the Chinese Academy of Sciences. H. Z. acknowledges the support of the International Postdoctoral Exchange Fellowship Program of China under Grant 20140029 and gratefully acknowledges the support of SA-SIBS scholarship program.

References


Chavez, A., Scheiman, J., Vora, S., Pruitt, B. W., Tuttle, M., E, P. R. I., Lin, S., Kiani, S., Guzman, C. D., Wiegand, D. J,


Li, T., Liu, B., Chen, C. Y., and Yang, B. 2016c. TALEN-mediated homologous recombination produces site-directed...


Sun, Y., Zhang, X., Wu, C., He, Y., Ma, Y., Hou, H., Guo, X., Du, W., Zhao, Y., and Xia, L. 2016b. Engineering herbicide-


