CellPress

Feature Review

Characteristics of Genome Editing Mutations in Cereal Crops

Changfu Zhu,[1,6] Luisa Bortesi,[2,6] Can Baysal,[1] Richard M. Twyman,[3] Rainer Fischer,[2,4] Teresa Capell,[1] Stefan Schillberg,[4] and Paul Christou[1,5,*]

Designer nucleases allow the creation of new plant genotypes by introducing precisely-targeted double-strand breaks that are resolved by endogenous repair pathways. The major nuclease technologies are meganucleases, zinc-finger nucleases, transcription activator-like effector nucleases, and the CRISPR/Cas9 system. Each comprises a promiscuous endonuclease guided by protein–DNA or RNA–DNA interactions. A great deal is known about the principles of designer nucleases but much remains to be learned about their detailed behavioral characteristics in different plant species. The outcome of genome engineering reflects the intrinsic properties of each nuclease and target genome, causing variations in efficiency, accuracy, and mutation structure. In this article, we critically discuss the activities of designer nucleases in different cereals representing a broad range of genome characteristics.

Genome Editing in Plants

The modification of plant genomes has been practiced since antiquity, initially through selective breeding, then more recently by mutagenesis and transgenesis, and in the last few years by genome editing [1]. The latter is a major advance because it does not depend on random recombination or integration events. Instead, double-strand breaks (DSBs) are introduced at precise sites in the plant genome by site-directed nucleases, allowing targeted modifications to be induced when the breaks are repaired by **non-homologous end joining (NHEJ)** or **homology-dependent repair (HDR)** [2,3]. The potential outcomes of genome editing are summarized in Figure 1. Genome editing was first achieved using natural **meganucleases** with target sites up to 18 bp in length [4] but the position of the DSB is determined by the natural specificity of the enzyme which is laborious to modify [5]. This limitation has been overcome by designing nucleases with bespoke specificity – the two principal examples are **zinc-finger endonucleases (ZFNs)**, in which a promiscuous endonuclease domain is paired with multiple zinc-finger DNA-binding domains each recognizing a 3-bp module [6], and **transcription activator-like effector nucleases (TALENs)**, in which a promiscuous endonuclease domain is paired with multiple transcription activator-like effector domains that recognize single base pairs [7]. The fourth and most recent genome editing technology is based on a form of bacterial adaptive immunity that neutralizes previously encountered invasive DNA sequences by expressing **clustered regularly interspaced short palindromic repeats (CRISPRs)** representing DNA fragments (**spacers**) captured from invading pathogens. The resulting CRISPR RNAs act as guides for CRISPR-associated (Cas) nucleases that attack the same pathogens if they enter the cell again [8]. Genome editing using the CRISPR system is achieved by constructing **synthetic guide RNAs (sgRNAs)** that direct the Cas nuclease to genomic targets, in contrast

Trends

Genome editing strategies based on designer nucleases (meganucleases, ZFNs, TALENs, and CRISPR/Cas9) have revolutionized plant breeding by allowing the introduction of insertions, deletions and substitutions at pre-determined sites.

The combination of genome editing and synthetic biology allows the creation of purpose-designed plants with precise sequence modifications.

The current body of data suggests that the CRISPR, ZFN, and TALEN methods are more efficient than meganucleases.

CRISPR (without enhancements) is more susceptible to off-target effects and great care is required during target selection to minimize the likelihood of unwanted mutations, particularly if the target has close paralogs in the genome.

The most effective way to generate homozygous plants is to use the most efficient nucleases, express them at high levels and (for HDR) deliver large amounts of donor DNA.

The outcome of both NHEJ and HDR editing seems to be consistent among the different methods, perhaps reflecting the overall similarity of the DSB repair mechanisms in cereals.

[1]Department of Plant Production and Forestry Science, School of Agrifood and Forestry Science and Engineering (ETSEA), University of

 http://dx.doi.org/10.1016/j.tplants.2016.08.009

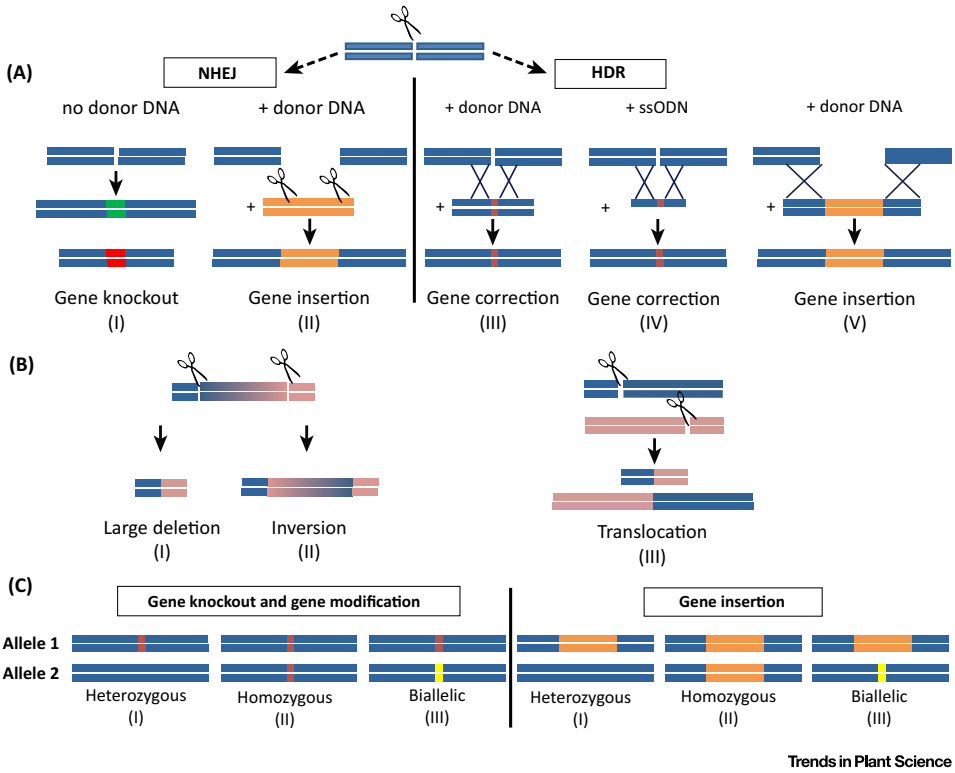


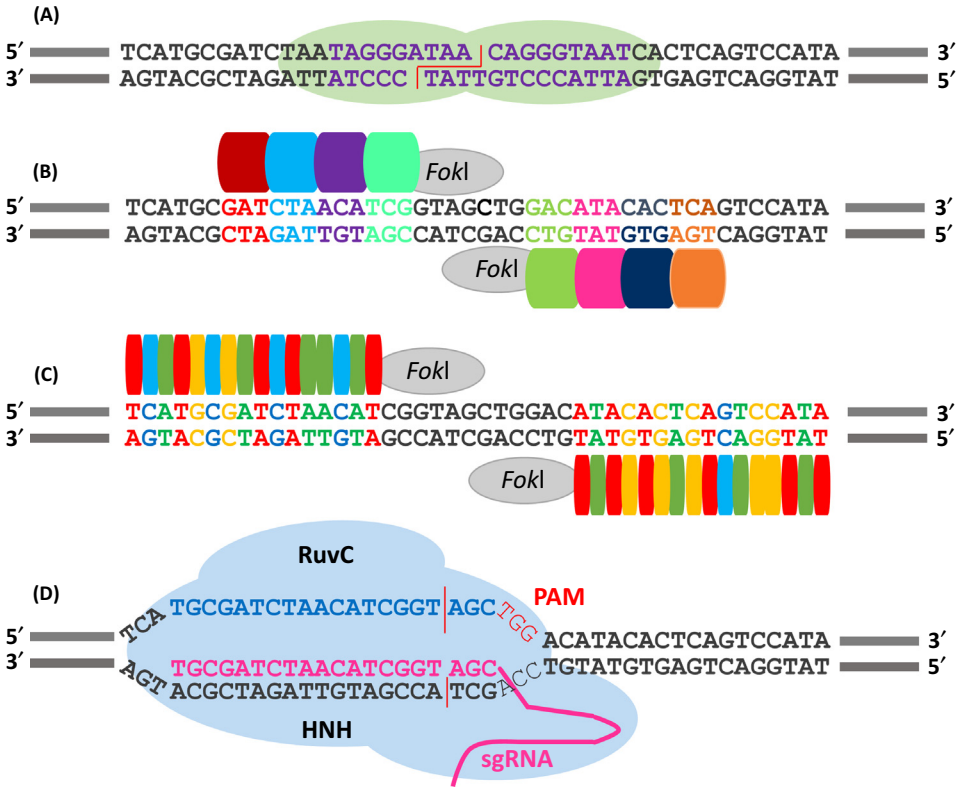
Figure 1. The Potential Outcomes of Genome Editing. Genome editing with designer nucleases can have multiple outcomes depending on the double-strand break (DSB) repair pathway and the nature of any donor DNA. (A) The major outcomes depending on the DSB repair pathway. The predominant repair pathway in plants is non-homologous end joining (NHEJ) which, in the absence of donor DNA (I), tends to generate short indels shown as insertions (green) or deletions (red) with all four types of nuclease. Both events tend to cause gene knockout. (II) If donor DNA is added to the cell which is flanked by the same target sites present in the genomic locus, then the designer nuclease will generate compatible ends and this can result in the integration of the new sequence, often accompanied by small indels at the junctions. Homology-dependent repair (HDR) occurs if a donor DNA template is available carrying the desired mutation. In the absence of donor DNA, HR involving sister chromatids will restore the locus to its original state (not shown). Donor DNA carrying a subtle change such as a nucleotide substitution can be provided as either a duplex molecule (III) or a single-stranded oligodeoxyribonucleotide (ssODN) (IV) and both will lead to allele correction. Alternatively, the homology region may be used to flank a new sequence (V), and this will lead to seamless gene insertion. (B) The possible outcomes when two DSBs are induced and repaired by NHEJ. If the DSBs are generated on a single chromosome the region between the two breaks can be deleted (I) or inverted (II). If the DSBs are induced on two different chromosomes, a chromosomal translocation can occur (III). (C) shows the potential consequences in diploid plants – the results of gene editing can be (I) heterozygous (single allelic change), (II) homozygous (identical changes to both alleles) or (III) biallelic (different changes at each allele) depending on which repair pathway is in operation. Adapted from [10].

to the other three systems which are protein-guided [9]. Genome editing is primarily based on the CRISPR/Cas9 system [10] although alternatives such as the CRISPR/Cpf1 system have been described more recently [11]. The four **designer nuclease** systems are compared in Figure 2.

Many articles have been published describing the use of designer nucleases for genome editing in different plants, as well as reviews that focus on targeting strategies, nuclease design, and applications for ZFNs [12], TALENs [13,14], and CRISPR/Cas9 [10,15–17]. These articles provide a rich source of data describing genome editing in different species, reflecting the unique combination of nuclease-specific properties (e.g., target site preference and DSB structure) and the characteristics of the target genome (e.g., size, GC content, and repetitive DNA content). Cereals provide a useful basis for such comparisons because genome editing has

Lleida-Agrotecnio Center, Lleida, Spain
[2]Institute for Molecular Biotechnology, RWTH Aachen University, Worringerweg 1, 52074 Aachen, Germany
[3]TRM Ltd, PO Box 463, York YO11 9FJ, United Kingdom
[4]Fraunhofer Institute for Molecular Biology and Applied Ecology IME, Forckenbeckstrasse 6, 52074 Aachen, Germany
[5]ICREA, Catalan Institute for Research and Advanced Studies, Passeig Lluís Companys 23, 08010 Barcelona, Spain
[6]These authors contributed equally to this work.

*Correspondence: christou@pvcf.udl.cat (P. Christou).



Trends in Plant Science

Figure 2. The Four Major Nuclease Technologies Used for Genome Editing. Genome editing is undertaken with four major classes of designer nucleases. (A) Meganucleases are usually natural endonucleases (or hybrids, or mutated/engineered variants) that recognize large targets, up to 18 bp in length, usually with twofold rotational symmetry. The DSB is made within the target site and is a staggered cut with overhangs (shown as a red line). Target site specificity therefore depends on the natural specificity of the enzyme and any variation that can be introduced by creating hybrid enzymes or selecting variant binding specificities by mutation and screening. This figure represents the natural meganuclease I-*Sce*I, which has the 18-bp recognition site 5′-TAG GGA TAA CAG GGT AAT-3′ and leaves four-nucleotide 3′ overhangs. (B) ZFNs are designer nucleases comprising the endonuclease domain from a restriction enzyme such as *Fok*I, where the DSB is introduced outside the recognition site, combined with a series of zinc-finger DNA-binding modules each with specificity for a 3-bp sequence. A bespoke nuclease can therefore be generated by arranging several zinc fingers as a tandem array; in this case four modules are used on each strand to create a 24-bp target with an intervening cleavage site, which *Fok*I cleaves nonspecifically leaving 4–5-nucleotide 5′ overhangs. Target specificity can be designed by selecting appropriate combinations of zinc fingers. (C) TALENs are similar in principle to ZFNs, featuring the same arrangement of endonuclease and DNA-binding modules. In this case, however, each TAL effector domain recognizes a single nucleotide pair, so the size of the target site reflects the number of TAL effector domains included in the nuclease. (D) The CRISPR/Cas9 system differs from the other three in that the specificity of binding is determined by a synthetic guide RNA (sgRNA). The sgRNA combines two functions of the natural CRISPR system, namely the specification of the target sequence, and the formation of a hairpin that helps to assemble the Cas9 nuclease complex. Cas9 can bind to any DNA sequence containing the relatively non-specific PAM site, but the sgRNA is required to identify the target for cleavage. Unlike the protein-guided nucleases, Cas9 introduces a blunt DSB three nucleotides upstream of the PAM so no single-strand overhangs are produced. However, the alternative CRISPR/Cpf1 system (not shown) works in an analogous manner but does introduce a staggered DSB. A similar outcome can be achieved by assembling two CRISPR/Cas9 complexes on distinct but adjacent targets and inactivating one of the nuclease domains so that each complex introduces a single-strand break (not shown). Abbreviations: DSB, double-strand break; PAM, protospacer adjacent motif; TALENs, TAL effector nucleases; ZFNs, zinc-finger nucleases.

been applied extensively to different cereal species, with multiple studies looking at the applications of genome editing in rice (*Oryza sativa*) and maize (*Zea mays*) as well as a smaller number of investigations in barley (*Hordeum vulgare*), wheat (*Triticum aestivum*), sorghum (*Sorghum bicolor*), and the model monocot *Brachypodium distachyon*. Cereal genomes differ greatly in size, complexity and sequence characteristics (Table 1), which allows us to draw comparisons

Glossary

Clustered regularly interspaced short palindromic repeats (CRISPR): bacterial DNA sequences containing repeats interspersed with spacer DNA from a previously encountered pathogen, such as a virus or plasmid.
***Cpf1*:** an RNA-guided endonuclease that does not need a separate tracrRNA, with similar applications to Cas9. It recognizes different PAMs compared to Cas9 and introduces a staggered DSB.
CRISPR-associated protein 9 (Cas9): an RNA-guided endonuclease whose natural function is to associate with crRNA and to degrade invasive DNA that has been encountered before, and whose major application is genome editing using sgRNA targeting specific genome sites.
CRISPR RNA (crRNA): short RNA cleaved from a long CRISPR transcript in bacteria. These associate with Cas9 and tracrRNA to form a targeted endonuclease complex that degrades invasive DNA.
Designer nuclease: an enzyme that cleaves nucleic acids, which has been selected or designed to cleave at a specific target site.
Double-strand break (DSB) repair: any endogenous pathway for the repair of double-stranded DNA breaks.
Genome editing: the introduction of precise changes at a specific locus in a genome.
HNH domain: one of two endonuclease domains in Cas9 – this cleaves the complementary strand.
Homologous recombination: a form of genetic recombination in which sequences are exchanged between two similar or identical molecules of DNA.
Homology-dependent repair (HDR): a DSB repair pathway that involves homologous recombination between two similar or identical sequences and therefore requires two homologous templates.
Indel: a general term for insertion or deletion.
Meganuclease: an endonuclease that recognizes a much larger target site (up to 18 bp) than a typical restriction endonuclease (4–8 bp) and therefore generates megabase fragments of DNA.
Nonhomologous end joining (NHEJ): a DSB repair pathway that

Table 1. Properties of cereal genomes. Data taken from [76–80]

Species	Genome size (Mb)	Ploidy	Gene number	Repetitive DNA (%)	GC content (%)
Barley	~5100	*n* = 7, 2n	79 379	Very high: 90	45%
Brachypodium	272	*n* = 10, 2n	42 868	Low: 21	45.9%
Maize	~2300	*n* = 10, 2n	63 540	High: 85	47.2%
Rice	403	*n* = 12, 2n	66 338	Low: 35	43.6%
Sorghum	818	*n* = 10, 2n	40 599	Low: 54	37.7%
Wheat	~17 000	*n* = 22, 2n = 6×	~96 000	Very high: 90	44.7%

between the species and to identify species-dependent and nuclease-dependent factors affecting the outcome of genome editing. This approach may facilitate the better design of genome editing strategies in the future.

Target Sequence Design

The sequences that can be targeted by native meganucleases are very limited. Different specificities can be introduced by generating hybrid or mutant variants but this process is challenging because the DNA-binding and endonuclease activities reside on the same domain [4,5]. In contrast, the DNA-binding and endonuclease components of ZFNs and TALENs are separate. In each case, the non-specific *Fok*I endonuclease domain is fused to a series of DNA-binding modules. The specificity of ZFNs and TALENs arises from the ability to customize these modules [18]. Most zinc-finger modules bind to a specific triplet when tested in isolation [19] but some show varying degrees of degeneracy or context-dependent promiscuity as part of a module, which means that reliable ZFN targets can be found approximately every 100 bp in genomic DNA [20]. TALENs are considered to bind with greater specificity than ZFNs although the individual modules differ in binding efficiency, and homopolymer runs should be avoided [21]. One limitation for the selection of TALEN targets is the requirement for thymidine at the first position [22]. Reliable TALEN cleavage sites can be found approximately every 35 bp in genomic DNA [23].

Unlike the other three systems, CRISPR/Cas9 specificity relies on RNA–DNA pairing at 20-nt genomic targets (described as 'spacers'). It is possible to find unique spacers in most genes, but in closely-related paralogs the unique sites tend to be more prevalent towards the 3′ end, particularly in the 3′ untranslated region [24]. However, unique sites are not usually necessary due to two further targeting constraints.

First, CRISPR/Cas9 targets must be preceded by a **protospacer adjacent motif (PAM)** which is necessary for Cas9 binding. The sgRNA/Cas9 complex recognizes two PAMs (5′-NGG-3′ and 5′-NAG-3′) but has a lower affinity and less tolerance for mismatches at the NAG-PAM [23,25]. The NGG-PAM is short and nonspecific and is therefore abundant in plant genomes, with one NGG-PAM found every ~10 bp in cereals; for example, 8.9, 9.8, 11.4, and 8.4 bp in brachypodium, rice, sorghum, and maize, respectively [26].

Second, the design of sgRNAs must also take into account the fact that Cas9 tolerates up to three mismatches in the sgRNA–DNA paired region, although the presence of mismatches near the NGG-PAM (the **seed region**) greatly reduces the affinity of Cas9 for the target site, so mismatches can only be tolerated at distal positions in the protospacer. This greatly increases the number of available protospacers. For example, a comparison of four cereal genomes shows that the total number of unique 20-nt protospacers is 13 743 in brachypodium, 11 677 in rice, 4982 in sorghum, and just 78 in maize, which would make maize an unsuitable CRISPR host if unique targets were necessary [26]. But taking into account that specific targeting can be

does not require extensive homology between templates and predominantly involves the direct ligation of broken ends, although short overlapping regions (microhomology) caused by staggered breaks do facilitate the repair process.

One-sided integration: a partially successful HDR in which one side of the cassette integrates by homologous recombination but the other by NHEJ.

Protospacer: the genomic target that is complemented by the sgRNA when CRISPR/Cas9 is used for genome editing.

Protospacer adjacent motif (PAM): a short and degenerate sequence (e.g., 5′-NGG-3′ in the case of Cas9) which is necessary for defining a target site for genome editing.

RuvC domain: one of two endonuclease domains in Cas9 – this cleaves the non-complementary strand.

Seed region: part of the spacer in a CRISPR sequence which is adjacent to the PAM and wherein base pairing is strictly required. The seed region of a sgRNA must therefore match the genomic target precisely.

Spacer: in the natural CRISPR system, spacers are the pieces of invasive DNA captured into the CRISPR sequences in the bacterial genome. The same term is sometimes adopted to also describe the sequence between the ZFN and TALEN binding sites that is cleaved by the *Fok*I endonuclease domain.

Synthetic guide RNA (sgRNA): a synthetic RNA that combines the natural functions of tracrRNA and crRNA to provide a two-component CRISPR/Cas9 system for genome editing.

TAL effector nuclease (TALEN): a designer nuclease that combines the endonuclease domain from a restriction enzyme such as *Fok*I with several transcription activator-like effectors (TALEs) which are DNA-binding proteins produced by phytopathogens to control plant genes in a way that favors bacterial infection. Combinations of TALEs can be used to form a bespoke DNA-binding component. Each TALE recognizes a single nucleotide pair and is surrounded by a TALEN scaffold that maintains the DNA-binding domains in the correct conformation.

achieved as long as there are four or more mismatches when aligned to any other 20-nt sequence, the number of available protospacers increases to 10 642 488 in brachypodium, 10 079 844 in rice, 11 046 150 in sorghum, and 10 180 095 in maize [26]. In these four cereals, the overall number of protospacers is thus comparable (10–11 $\times$ 10^6) and not proportional to the genome size, the total transcript number or the NGG-PAM number. This suggests that specific sgRNAs can be designed for 98.8% of annotated transcription units in brachypodium, 89.6% in rice, 93.9% in sorghum, and 30.5% in maize [26,27]. In some cases, sgRNAs that target multiple related genes may be desirable, for example, to knock out all members of a functionally redundant gene family [28].

Trans-acting CRISPR RNA (tracrRNA): a bacterial noncoding RNA that binds to crRNA and forms a complex with Cas9.

Zinc-finger nuclease: a designer nuclease that combines the endonuclease domain from a restriction enzyme such as *FokI* with several zinc-coordinating DNA-binding domains (usually derived from a transcription factor) that form a bespoke DNA-binding component. Each zinc-finger domain recognizes a 3-bp DNA sequence

Genome Editing Using the NHEJ Pathway

Single Cut

Most genome editing studies in plants involve the introduction of one targeted DSB which is repaired by the erroneous NHEJ pathway, resulting in the formation of an **indel** or in some cases a substitution at the site of the break. The efficiency of genome editing can be defined as a percentage, representing the number of positive targeting events (where indels are formed) per 100 transformations. This is sometimes described in the literature as the mutation frequency. The accuracy can be defined as the number of events that have the desired outcome compared to the number of off-target mutations (indels at unintended sites). It is also useful to compare the structure of the indels (size, relative prevalence of insertions and deletions, representation of different nucleotides). The data from a large number of published studies involving cereal genomes edited with the four nuclease systems are summarized in Table 2.

In rice, mutation frequencies in the range 4–30% have been reported in callus and T0 plants using TALENs [29–34]. The T0 plants generated by TALENs are often chimeric, but heterozygous, biallelic, and homozygous mutations at three targets have been reported [32,33]. Most genome editing studies in rice have involved the CRISPR/Cas9 system, and much higher mutation frequencies have been achieved using this approach compared to TALENs [35–44]. The average mutation frequency induced by CRISPR/Cas9 in rice is 85% when comparing 46 target sites [31] and in some cases the mutation frequency in T0 plants has reached 100% [37,39]. Biallelic mutants have been reported frequently in the T0 generation [40–43,45] and represented up to 100% of the mutations when the most efficient sgRNA was used [39]. Furthermore, homozygous plants were obtained at frequencies of up to ~50% in the T0 generation [31,46,47]. The use of multiple sgRNAs targeting the same gene at different locations further increased the frequency of homozygous mutants [47]. CRISPR/Cas9 efficiently generated mutations in three paralogous rice genes using one sgRNA, taking advantage of the tolerance of the system to mismatches at distal positions in the protospacer [45]. Nearly all mutations induced by TALENs and CRISPR/Cas9 in rice are indels, typically 1-bp insertions or

Box 1. Detailed Analysis of Targeted Mutations

The presence of target mutations can easily be determined using a mismatch-cleavage assay with T7EI or Surveyor™ nuclease [83], which cleave heteroduplex DNA at mismatches and extrahelical loops formed by single or multiple unpaired nucleotides [84]. Alternatively, the digestion of PCR products with restriction enzymes is suitable when the designer nuclease cuts within or adjacent to a restriction enzyme recognition site. The size of indels can be determined by high-resolution melting analysis (HRMA), which can also detect homozygous mutations [85]. The gold standard for the analysis of target mutations is the cloning and sequencing of individual amplicons. Bulk PCR products can also be analyzed by Sanger sequencing and subsequent decoding of superimposed chromatograms [86,87] or by next generation sequencing (NGS) followed by software-based analysis [88]. Gene integration events can be identified by sequencing, junctional 'in-out' PCR [70], and Southern blot analysis. The analytical procedures used to detect on-target mutations can also be applied to mutations at off-target sites that are similar to the target sequence [89]. These can be predicted using several online bioinformatics tools (see examples[i,ii,iii] in Resources) and, as discussed earlier, steps can be taken to design targeting strategies that reduce the likelihood of mutations at off-target sites even if the genes are closely related. Unbiased methods to detect unintended mutations include whole exome or whole genome sequencing, and procedures that track DSBs at any genomic sites, including BLESS [90] and GUIDE-seq [64].

Table 2. The efficiency and accuracy of designer nucleases in different cereals

Tissue[a]	Input[b]	Repair[c]	Percentage efficiency and accuracy of editing[d]								Type and size of mutations[e] (bp)				Comments	Refs
			MN		ZFN		TALENs		CRISPR							
			ON	OFF	ON	OFF	ON	OFF	ON	OFF	DEL	INS	SUB	COM		
Barley																
T0	A	NHEJ	–	–	–	–	22	NF			4–36	10	–	X	–	[56]
T0	A	NHEJ	–	–	–	–	–	–	10–23	4.2	1–3	1	–	X	–	[57]
Brachypodium																
P	A	NHEJ	–	–	–	–	8–10	NA			≤74	≤78	–	X	–	[81]
Maize																
E	A	NHEJ	1	NA	–	–	–	–	–	–	≤60	1–17		X	–	[48]
E	A	NHEJ	3	NF	–	–	–	–	–	–	≤220	1–12	–	X	Monoallelic, biallelic, and chimeric mutations identified at the *liguleless1* locus. Approximately half (10/23) of the mutant T0 plants contained monoallelic or biallelic mutations at the *lig34* locus	[49]
T0	A	NHEJ	5.8	NA	–	–	–	–	–	–	1	1	–	–	Monoallelic and biallelic mutations identified	[50]
C	W	HDR	–	–	19–40	NF	–	–	–	–	2–7	4	–	–	Efficiency of 19% using an autonomous selection marker but 40% using a nonautonomous marker	[60]
E	A	NHEJ	–	–	–	–	3.7–10	NF	–	–	3–9	–	–	–	Efficiency of 3.7% (1/27) for the inbred line and 9.9% (9/91) for hybrid T0 plants	[51]
T0	A	NHEJ	–	–	–	–	13–39	NA	–	–	2–10	–	–	–		[52]
P			–	–	–	–	9.1–23.1	NA	–	–	2–27	–	–	–		
P			–	–	–	–			16.4–19.1	NA	≤44	125		X		
Rice																
T1	A	NHEJ	–	–	–	–	48–63	NF			≤55	3–9		X		[29]
T0	P	NHEJ	–	–	–	–	30	NF			≤72	≤124	–	X	Three potential off-target sites were analyzed with 9-bp (OffT-1), 11-bp (OffT-5), and 10-bp (OffT-17) mismatches compared to the target. No off-target mutations found in the T0 plants	[32]
T0	A	NHEJ	–	–	–	–	4	NF			4–10	1	–	–	–	[31]
T0	A	NHEJ	–	–	–	–	–	–	21–67	10	≤100	1	1	X	Low-frequency mutations were found in only one off-target site among 13, involving a 1-bp difference from the intended target outside the seed region	[35]
P	P	HDR	–	–	–	–	–	–	7.1–9.4	NF		12	–	–	No evidence of sgRNA/Cas9-induced off-target mutations was found using the PCR/restriction endonuclease assay (Box 1)	[36]
T0		NHEJ	–	–	–	–	–	–		NF	1–207	24	–	–		
T0	A	NHEJ	–	–	–	–	–	–	27.5–67.5	2.5–67	≤85	1–2	–	–	Accuracy differed when targeting the *Gn1a* and *DEP1* genes depending on the number of mismatches in the sgRNA (see main text for details)	[44]

Table 2. (continued)

Tissue[a]	Input[b]	Repair[c]	Percentage efficiency and accuracy of editing[d]								Type and size of mutations[e] (bp)				Comments	Refs
			MN		ZFN		TALENs		CRISPR							
			ON	OFF	ON	OFF	ON	OFF	ON	OFF	DEL	INS	SUB	COM		
TO	A	NHEJ	–	–	–	–	–	–	4–29	13–22	≤33	1	–	–	No mutations found at off-target sites with mismatches 6, 14, and 17 bp from the PAM (10th most likely off-target according to CRISPR-P) and also no mutations in 3rd, 5th, and 9th potential off-targets. Biallelic mutation in off-target gene but not target gene	[45]
P	A	NHEJ	–	–	–	–	–	–	3–8	1.6	3–14	42–195	–	X	–	[27]
T0	A	NHEJ	–	–	–	–	–	–	4.8–75	NF	1–16	1	–		–	[59]
T0	A	NHEJ	–	–	–	–	–	–	2.1– 15.6	NF	1–4	5	1	X	Three highly conserved sites were identified with mismatches of 1 or 3 bp compared to the target, but no off-target mutations were found	[40]
Sorghum																
E	A	NHEJ	–	–	–	–	–	–	30	NF	1–9	1	–	–	–	[82]
Wheat																
T0	P	NHEJ	–	–	–	–	3.4–6	NA			≤29	≤141	–	X	The mutation frequency for *TaMLO-A1* in T0 plants was also tested by CRISPR/Cas9 and similar results were obtained (5.6%) compared to TALENs	[28]
P			–	–	–	–	23–38	NA	–	–	1–32	61–141	–	X		
P	P	NHEJ	–	–	–	–	–	–	28.5	NF	2–10		–			[36]
S	A	NHEJ	–	–	–	–	–	–	18–22	1–3	≤53	≤22	–		Duplex sgRNA/Cas9 achieved 2.8% efficiency for 53-bp deletions between sites	[24]

[a]Abbreviations: C, callus; E, embryos; P, protoplasts; S, suspension culture; T0, T0 plants; T1, T1 plants.
[b]Abbreviations: A, Agrobacterium; P, particle bombardment; W, whisker-mediated.
[c]Abbreviations: HDR, homology-dependent repair; NHEJ, non-homologous end joining.
[d]In these columns; efficiency and accuracy are defined by the percentage of on-target and off-target mutations. Efficiency = (on-target + off-target)/total attempts. Accuracy = on-target/total attempts. Abbreviations: CRISPR, CRISPR/Cas9 system; NA, no analysis was carried out by the authors; NF, analysis was carried out by the authors but no mutations of this category were found; MN, meganuclease; TALEN, TAL effector nuclease; ZFN, zinc-finger nuclease.
[e]An X in column COM shows that combinations of DEL, INS and/or SUB events were recovered. Abbreviations: COM, combination; DEL, deletion; INS, insertion; SUB, substitution.

deletions ranging from one to hundreds of base pairs, whereas only 4% of the reported mutations were single-nucleotide replacements [29–44].

In maize, NHEJ efficiencies of 1–5.8% were reported using meganucleases [48–50], and efficiencies of 10% [51] and 39% [52] up to 100% [53] have been reported using TALENs. When directly compared, the CRISPR/Cas9 and TALEN approaches achieved similar mutation frequencies at the same target site (13.1% vs. 9.1%), and CRISPR/Cas9 achieved up to a 20-fold higher mutation frequency than meganucleases [53]. For all three nucleases, biallelic mutants have readily been recovered in the T0 generation [49–51,53]. The size and efficiency of deletions varies by method; for example, 3–10 bp with up to 10% efficiency in T0 plants by TALENs [51], and 2–220 bp with 1–3% efficiency centered at the cleavage site using meganucleases [48,52]. CRISPR/Cas9 and TALENs both generated indels with 9–13% efficiency in maize [52].

In barley, the efficiency of insertions was 2.1% with **one-sided integration** when using meganucleases [54]. Three TALEN sites were selected based on their proximity to the target region and two different Talen-3 scaffolds were generated because of their increased cleavage activity: both TALEN constructs were able to bind to the specific target and achieved a 16% mutation frequency for the first construct and 31% for the second [55]. The use of TALENs in haploid cells achieved a mutation frequency of up to 22% [56], which was similar to the 23% in T0 plants reported using CRISPR/Cas9 [57]. In the latter case, CRISPR/Cas9 was used to target two copies of the same gene (90% sequence identity), achieving a 23% mutation frequency in the T0 generation for one copy and a 10% mutation frequency for the other, predominantly a mixture of 1-bp insertions and 1–36 bp deletions [57].

In hexaploid bread wheat, the TALEN and CRISPR/Cas9 systems were used to introduce targeted mutations in the three homoeoalleles that encode MILDEW-RESISTANCE LOCUS (MLO) proteins [28]. The TALEN approach yielded 27 mutations among 450 independent T0 transgenic plants (3.4% in spring wheat and 6.0% in winter wheat), and the majority of the mutants were heterozygous, with one homozygous and one heterozygous mutant for three homoeoalleles. The CRISPR/Cas9 system was then used to generate plants with mutations in a single allele. Four independent mutants were identified among 72 T0 transgenic wheat lines, achieving a mutation frequency of 5.6%. All except one homozygous mutant plant was obtained by selfing. The TALEN and CRISPR/Cas9 approaches were therefore comparable in efficiency at these loci but the CRISPR approach achieved greater targeting selectivity.

Double Cut

In a few studies, two DSBs have been introduced simultaneously in cereal genomes with the intention of deleting the intervening sequence. The frequency of targeted deletions between two DSBs depends on the efficiency of cleavage at each target. For the CRISPR/Cas9 system, sgRNAs with a higher GC content achieved greater editing efficiency in rice [31,35,40]. In wheat, a deletion of 53 bp was observed with a frequency of 2.8%, and mutations at each site were observed with a frequency of 11–12% [24]. In rice higher deletion frequencies were achieved using TALENs, that is, the deletion of a 1.3-kb fragment was achieved in 5.1% of callus clones [32]. Using the CRISPR/Cas9 approach, deletions of 200 bp were achieved at a frequency of 10% in rice protoplasts [38], and deletions of 357–761 bp were achieved at frequencies of 4–45% in rice protoplasts and up to 6% in T0 plants [37]. The deletion of large chromosomal segments (115–250 kb) has also been achieved at high frequencies of 16.7–25% in T0 rice plants using the CRISPR/Cas9 system [39]. If donor DNA is provided, a double cut can also be used to insert a gene cassette (Figure 1) although this has not been reported in cereals. Another possible outcome is a chromosomal inversion (Figure 1). The inversion of a 1.3-kb fragment was achieved with a frequency of 1% in rice using TALENs (compared to 5.1% deletions) [32], whereas the inversion of a larger chromosomal fragment was reported as a rare alternative to deletion in rice targeted with two sgRNA/Cas9 constructs [58].

Overall Principles

The differences in mutation signatures generated by genome engineering in major cereal crops appear to depend more on the nuclease than the species, possibly reflecting intrinsic similarities in the DNA repair pathways among monocotyledonous plants (Figure 3). All four nuclease types predominantly create indels rather than substitutions, and the difference in indel length mirrors the fact that Cas9 generates a blunt-ended DSB whereas the other three produce overhangs that can be trimmed during the NHEJ repair process.

The most common mutations induced by Cas9 are very short indels ($\leq$10 bp), often single nucleotides, and the insertions are mostly A/T base pairs as shown in barley [57], rice [31,35,37,41,42,44,47,59], and maize [53]. One report based on wheat cell suspension cultures

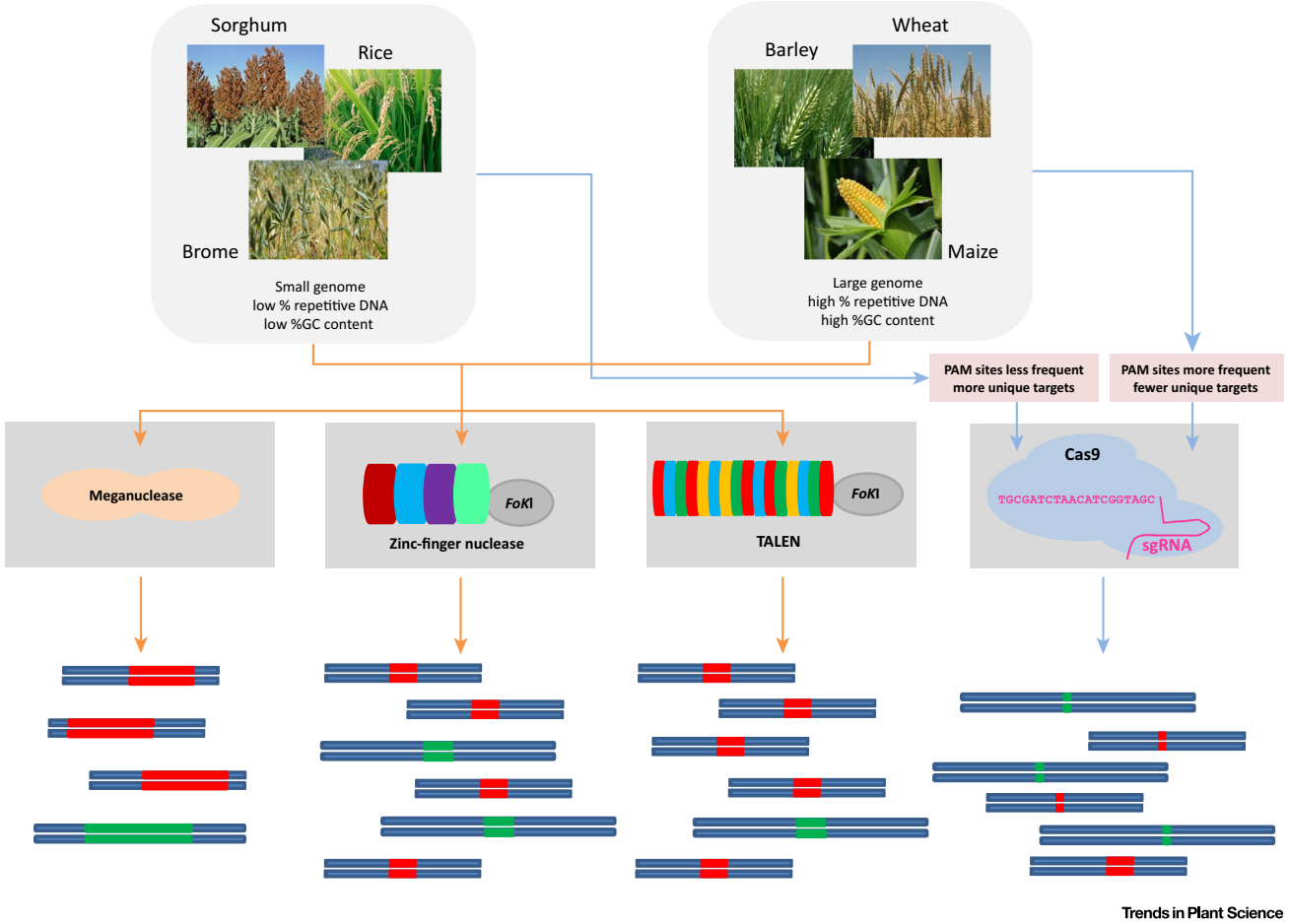


Figure 3. Structures of On-target and Off-target Mutations Caused by Four Different Types of Designer Nucleases in Different Cereal Species. It is assumed no donor DNA is provided and the predominant NHEJ repair pathway is induced. The outcome of genome editing depends on the properties of the target genome and the nuclease. In cereals with small genomes (rice, sorghum, and brachypodium), the low GC content and low proportion of repetitive DNA increases the number of unique targets for the CRISPR/Cas9 system because PAM sites are less frequent. In contrast, cereals with larger genomes (maize, barley and wheat) have more frequent PAM sites and fewer unique targets. Even so, because sgRNAs are designed with care to target unique sequences wherever possible, the outcome is similar in all cereals, comprising an approximately equal mixture of small (usually 1-bp, but occasionally larger) insertions (green) and deletions (red). The CRISPR/Cas9 pathway is shown with blue arrows. The other three classes of designer nuclease have equivalent effects in all cereals because they target unique exonic sequences (orange arrows). However, there are important differences in the outcome of genome editing with each class of nucleases. TALENs and ZFNs both contain the same *FokI* nuclease domain and induce DSBs with 5′ cohesive ends. TALENs tend to generate predominantly small deletions and occasionally insertions. ZFNs also produce predominantly small deletions, but a larger proportion of insertions than TALENs. This is probably due to the fact that the shorter spacers in the ZFN sites (mostly 5–6 bp) compared to the TALEN sites (12–21 bp) produce defined 4–5-bp overhangs, which can be efficiently filled in before ligation yielding insertions [69]. In both cases, the indels are typically a few bp in length and therefore generally longer than those introduced by CRISPR/Cas9, which generates a blunt DSB. In contrast, meganucleases generate short 3′ overhangs which predominantly result in larger deletions, with occasional larger insertions. The CRISPR/Cas9, TALEN, and ZFN editing pathways can also generate larger insertions and deletions, but these are rarer outcomes and predominantly reflect gene-specific effects. Abbreviations: NHEJ, nonhomologous end joining; PAM, protospacer adjacent motif; TALENs, TAL effector nucleases; ZFNs, zinc-finger nucleases.

described medium-size indels in one gene [24]. Further data are therefore required to determine whether this is an intrinsic feature of the wheat **double-strand break (DSB) repair** pathway, a property specific to the target gene, or a characteristic of the cell suspension culture. In reference to the latter possibility, larger indels were also observed in maize protoplasts engineered using CRISPR/Cas9, suggesting the cell status may play a role in the outcome of editing [52].

In contrast to CRISPR/Cas9, meganucleases tend to generate deletions, which range in size from 2 to 71 bp in maize [48,49]. Even when the same gene is targeted, the peculiarities of the CRISPR/Cas9 and meganuclease events are preserved, suggesting an intrinsic difference in

the nature of the DSBs and/or their repair [53]. TALENs appear to induce predominantly multiple base-pair deletions and only a few short insertions or combinations of insertions and deletions as reported in barley [56], maize [51], rice [29,30,32,36], and wheat [28]. The only work describing ZFN-induced mutations in maize indicated the predominant occurrence of short deletions, 2–7 bp in length [60]. ZFNs tend to be associated with more insertions than TALENs, and this was proposed as a potential disadvantage of the method if engineered cereals are taken through the regulatory approval process [61].

Although it is possible to identify trends in mutation structures for each type of nuclease, all four can also generate larger indels of 100–300 bp. Any treatment or physiological state that affects the DNA repair pathways can influence the type of mutations recovered; for example, rice mutants lacking the NHEJ enzyme DNA ligase 4 (Lig4) generate longer indels at TALEN target sites [33]. Some differences can be attributed to variable experimental conditions, but different mutation signatures have been reported within the same species at different loci. For example, CRISPR/Cas9 was shown to predominantly introduce large deletions of 94–760 bp in some rice genes [46], whereas others almost exclusively featured small deletions of $\leq$11 bp [39] or 1-bp insertions [35]. In the same study, two genes predominantly featured short indels of 1–8 bp while a third gene featured only longer deletions of 11–44 bp [43]. The editing of maize genes with meganucleases [49] and rice genes using TALENs [33] revealed the frequent presence of microhomology regions at the ends of the largest deletions. These observations indicate that gene-specific factors can influence the outcome of DSB repair, regardless of the nuclease type.

Detailed Analysis of Off-target Mutations

ZFNs and TALENs tend not to cleave at off-target sites because they have long recognition sequences that must be a certain distance from the cleavage position. Therefore, off-target sites are mostly an issue with the CRISPR/Cas9 system because it can tolerate mismatches between the sgRNA and the target, and most investigations of off-target mutations have therefore focused on this method. Mismatches within the first 12 bp proximal to the PAM generally abolish DNA cleavage [62] and this has indeed been observed in rice [27,34,36,39,40,43], wheat [24], and maize [63] even when large-scale effects in rice were examined by whole-genome sequencing [35]. However, 47.5% off-target mutations were observed in rice at a site with one mismatch at position 8 upstream from the PAM, using a sgRNA with a 65% GC content targeting the *IPA1* locus [44]. In contrast, only 2.5% off-target mutations were observed using a sgRNA targeting *DEP1*, which differed at six positions from the off-target site.

Furthermore, the induction of off-target mutations has also been exploited intentionally to achieve multi-gene knockout in rice expressing one sgRNA. In this case, the overall sgRNA GC content was 67% and the tolerated mismatch was an A replacing a G at position 10 in the sequence CGGGGAGGGG [45]. Xie and Yang [27] observed mutations at an off-target site with one base deletion at position 11 in the middle of the sequence GGCGTGCCT within a sgRNA with an overall GC content of 59%. In barley, 4.2% off-target activity was detected in the T1 progeny of the line with the highest on-target activity, at a site with one mismatch at position 9 upstream from the PAM. In this case, the sgRNA had an overall GC content of 60%, and the tolerated mismatch was a C replacing T in the sequence CTGGGGC [57]. Although sgRNAs with a GC content greater than 70% may encourage off-target mutations [64], it is tempting to speculate that short GC stretches around the mismatch may compensate for the imperfect pairing of the sgRNA and target, but further studies are required to confirm this hypothesis. In barley lines with lower on-target activity [57], the T1 progeny did not contain indels at the off-target site. This suggests that high on-target efficiency, possibly due to the persistence of the CRISPR components following stable integration, may also promote cleavage at unwanted sites. Accordingly, off-target rice mutations were only observed in T1 plants containing the integrated *cas9/sgRNA* transgene [43]. The GC content of the protospacer therefore appears to

affect the efficiency and accuracy of CRISPR/Cas9 and this could restrict the number of targets in GC-rich genomes, although the large number of potential protospacers means that careful selection of targets with a balanced GC-content should overcome this limitation.

The high frequency of biallelic and homozygous T0 rice mutants recovered in many studies indicates that cleavage predominantly takes place in the transformed callus cells before regeneration. The transient expression of *cas9* and *sgRNA* genes (or any of the other nucleases) should hence suffice to induce on-target mutations at acceptable rates, while helping to avoid off-target cleavage. Overall, these data indicate that CRISPR/Cas9 tends to be highly specific in rice [35] and that the rare off-target mutations could be almost entirely eliminated by the careful selection of sgRNA sequences and shorter regeneration times. Interestingly, all off-targets identified in rice [44] were 1-bp insertions, suggesting that the pairing of sgRNA with the target sequence may also influence the mutation type. Recently, the use of Cas9 paired nickases resulted in the complete suppression of off-target mutations in rice callus and regenerated plants [65]. Moreover, engineered variants of Cas9 with greatly improved specificity have been tested in human cells and showed no detectable genome-wide off-target effects and no significant reduction of on-target activity [66,67].

Off-target mutations were not detected in the genome of T0 maize plants edited using ZFNs [60]. Furthermore, the longer recognition sites of TALENs are thought to induce even fewer off-target events than ZFNs and the CRISPR/Cas9 system [31]. The likelihood of off-target mutations in plant genomes edited using designer nucleases should be considered from a holistic perspective, because such unwanted mutations may be less frequent than the somatic mutations that occur randomly during tissue culture-based transformation [31].

Genome Editing Using the HDR Pathway

NHEJ is the prevalent DNA repair mechanism in eukaryotes, and even when DSBs are introduced, the frequency of targeted integration by HDR remains 10^3–10^6 times lower than random integration [68]. HDR-mediated gene targeting in cereals has been developed predominantly in maize, and remarkable gene targeting frequencies have been reported. Using the meganuclease I-*Sce*I, gene targeting was achieved at frequencies up to 7.6 events per bombarded plate, with a calculated homologous versus nonhomologous recombination ratio of up to 30%, and with up to 40% of the recovered events bearing no traces of random insertions [69]. I-*Sce*I has also been used for gene targeting in barley, where one-sided integration of the donor was observed in 2.1% (2/95) of the regenerated plants [54]. HDR directed by ZFNs produced small insertions with an efficiency of 20% using an autonomous *PAT* selection marker, and 40% using a nonautonomous *PAT* marker that required precise integration adjacent to the endogenous *IPK1* promoter for expression [60], although in another study frequencies of up to 5.3% were reported, along with position effects affecting targeting efficiency in two target lines [70]. A direct comparison of meganucleases and CRISPR/Cas9 in maize showed that the frequency of HDR events induced by CRISPR/Cas9 (up to 4%) was up to fivefold greater than the number induced by meganucleases (up to 0.7%) [53].

The number of studies involving HDR in rice is small compared to those involving NHEJ. The initially poor efficiency of **homologous recombination** in rice was increased by altering the equilibrium between the competing HDR and NHEJ pathways. For example, overexpressing the exonuclease OsExo1 and the helicase OsRecQl4 increased the efficiency of HDR by ~30-fold when using the meganuclease I-*Sce*I [71], and gene targeting frequencies of up to 1% could be achieved in rice Lig4 mutants by transforming callus with the appropriate sgRNA and homologous donor DNA to mutate the *ALS* gene [72]. Impressive results were recently achieved with an optimized strategy to introduce multiple discrete point mutations in the *ALS* gene [73]. The authors used two sgRNAs instead of one, flanked the homology arms on the donor plasmid with

CRISPR sites to release the repair template *in vivo*, and increased the amount of donor DNA by delivering the plasmid together with free donor fragments using particle bombardment. They found that 100% of T0 plants resistant to both selection markers had undergone HDR, including 92% (48/52) that were homozygous for the replacement. This is by far the best result ever reported in higher plants. Interestingly, when the same authors used single-stranded oligonucleotides as a repair template in rice (following their successful deployment in maize [53]) they did not recover any edited *ALS* events. These data suggest that different species may require distinct strategies to benefit from oligonucleotide-mediated repair, and it will be interesting to see whether the improved CRISPR-based HDR approach is equally successful in other species.

Particle bombardment appears to be more effective than *Agrobacterium tumefaciens* for the promotion of homologous recombination induced by meganucleases in maize [69] and CRISPR in maize and rice [53,73], not only increasing the efficiency of homologous recombination overall but also the ratio of HDR:NHEJ events [69]. This may reflect differences in the molecular environment created by particle bombardment and *Agrobacterium*-mediated transformation, with the former reported to introduce larger amounts of donor DNA that may induce a multi-tiered DNA repair response in which DNA repair complexes are recruited to sites of transgene integration, stimulating further repair and integration events [74]. The introduction of donor DNA by particle bombardment not only increased the efficiency of gene targeting in rice but also achieved 92% homozygous replacements, whereas *Agrobacterium*-mediated transformation produced 100% heterozygous HDR events with the opposing allele showing the footprints of NHEJ [73]. Increasing the amount of donor DNA can increase the likelihood of unwanted random integration events, although negative selection markers such as *cod*A placed outside the homology region can reduce the number of events that survive [75]. Interestingly, a twofold difference in gene targeting frequency was observed using ZFNs and the same artificial landing pad in two independent transgenic maize lines, suggesting that position effects like the epigenetic status of the locus and the genomic context of the target site can also greatly affect the experimental outcome [70].

Concluding Remarks

The availability of genome editing technologies based on four different types of nucleases with distinct mechanisms for the specification of target sites and the introduction of DSBs raises the possibility that they will show species-dependent effects due to differences in the general properties of the genome in each species. Meganucleases, ZFNs and TALENs each use protein-based recognition to identify target DNA sequences and have a separate nuclease domain that introduces a staggered DSB. There is a greater choice of target sites for ZFNs and TALENs because the DNA-binding component is constructed from modular units recognizing 3-bp (ZFN) or 1-bp (TALEN) subsequences, whereas meganucleases recognize longer sites that are dependent on natural specificity and any variations that can be achieved by mutagenesis. In contrast, Cas9 forms a complex with a sgRNA to locate its target sequence (although the protein also binds DNA directly at the PAM site) and introduces a blunt DSB. The more recent CRISPR/Cpf1 system recognizes a distinct T-rich PAM site (which reduces the likelihood of off-target sites in GC-rich genomes although also reducing the number of unique sites overall) and the DSB is staggered. The nature of the DSB and the presence or absence of donor DNA plays a key role in determining the balance between NHEJ and HDR, and thus the outcome of genome editing, so variations of the CRISPR/Cas9 system in which either the **HNH** or **RuvC endonuclease domains** are disabled, or two Cas9-sgRNA complexes target adjacent sites, can therefore be used to encourage the introduction of specific types of mutations.

The current body of data suggests that the CRISPR, ZFN, and TALEN methods are more efficient than meganucleases, but that CRISPR (without enhancements) is more susceptible to

Outstanding Questions

The strict identification of species-dependent differences between cereals would require the same gene to be targeted for editing in different species using a comparable technique. This may be possible with a highly conserved gene, or perhaps with an identical transgene introduced into different species specifically for the purpose of nuclease analysis.

It would also be useful to identify specific components of the DNA repair pathways that influence the outcome of editing because these may provide the most effective route to increase the efficiency and accuracy of targeted mutations in plants, and the ability to introduce precise changes on demand.

off-target effects and great care is required during target selection to minimize the likelihood of unwanted mutations, particularly if the target has close paralogs in the genome. The most effective way to generate homozygous plants is to use the most efficient nucleases (and sgRNAs with a higher GC content in the CRISPR system), to express nucleases at high levels and deliver large amounts of donor DNA for HDR, and in at least rice and maize to deliver the components by direct DNA transfer rather than *Agrobacterium*-mediated transformation. Multiple sgRNAs can also increase efficiency [28].

Nuclease-dependent differences in mutation structure have been identified (Figure 3). For example, whereas Cas9 tends to generate short indels (1–10 bp) with an equal likelihood of insertions or deletions, TALENs and ZFNs are more likely to produce deletions, but ZFNs yield more insertions than TALENs, perhaps due to the difference in the size of the spacers, which are typically 12–21 bp in the case of TALENs and 5–6 bp in the case of ZFNs [61]. Meganucleases are more likely to cause larger deletions (~70 bp) but all four methods can yield occasional larger indels. These differences between methods are somewhat overshadowed by variability caused by the targeted gene locus and the physiological state of the tissue, with undifferentiated tissue and suspension cells more likely to show more frequent and extensive mutations. Species-dependent effects are more difficult to identify. Despite the significant differences in size, repetitive DNA content and GC content, for example between the genomes of rice (small, little highly-repetitive DNA, relatively low GC content) and maize (large, abundant highly-repetitive DNA, high GC content), the outcome of both NHEJ and HDR following genome editing seems to be consistent among the different methods. Perhaps this reflects the overall similarity of the DSB repair mechanisms in cereals and the tendency, despite differences in genome composition, to target sequences with similar characteristics for genome editing (i.e., exons in genes, typically unique sequences that tend to have a more consistent GC content across genomes than the average values, which are skewed by the properties of repetitive DNA). Nevertheless, some interesting exceptions have been observed, such as the failure of single-stranded repair templates to promote genome editing in rice but their ability to improve the efficiency of gene targeting in maize [53,73] (see also Outstanding Questions).

Acknowledgments

Synthetic biology and genome editing work at the UdL is supported by grants from the Spanish Ministry of Economy and Competitiveness (BIO2014-54426-P and BIO2014-54441-P) and the Catalan Government 2014 SGR 1296 Agricultural Biotechnology Research Group. Work on gene targeting at the RWTH Aachen University is funded by the European Research Council Advanced Grant 'Future-Pharma', Grant Number 269110. Fraunhofer IME has received funding from Dow AgroSciences for research on zinc-finger nucleases.

Resources

[i] http://bao.rice.edu/cgi-bin/prognos/prognos.cgi

[ii] www.e-talen.org/E-TALEN/designTALENs.html

[iii] www.e-crisp.org/E-CRISP/

References

1. Weeks, D.P. *et al.* (2016) Use of designer nucleases for targeted gene and genome editing in plants. *Plant Biotechnol. J.* 14, 483–495
2. Baltes, N.J. and Voytas, D.F. (2015) Enabling plant synthetic biology through genome engineering. *Trends Biotechnol.* 33, 120–131
3. Osakabe, Y. and Osakabe, K. (2015) Genome editing with engineered nucleases in plants. *Plant Cell Physiol.* 56, 389–400
4. Smith, J. *et al.* (2006) A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. *Nucleic Acids Res.* 34, e149
5. Rosen, L.E. *et al.* (2006) Homing endonuclease I-CreI derivatives with novel DNA target specificities. *Nucleic Acids Res.* 34, 4791–4800
6. Kim, Y.G. *et al.* (1996) Hybrid restriction enzymes: zinc finger fusions to FokI cleavage domain. *Proc. Natl. Acad. Sci. U.S.A.* 93, 1156–1160
7. Christian, M. *et al.* (2010) Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186, 757–761
8. Marraffini, L.A. and Sontheimer, E.J. (2010) CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat. Rev. Genet.* 11, 181–190
9. Jinek, M. *et al.* (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821
10. Bortesi, L. and Fischer, R. (2015) The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnol. Adv.* 33, 41–52

11. Zetsche, B. *et al.* (2015) *Cpf1* is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163, 759–771
12. Petolino, J.F. (2015) Genome editing in plants via designed zinc finger nucleases. *In Vitro Cell. Dev. Biol. Plant* 51, 1–8
13. Doyle, E.L. *et al.* (2013) TAL effectors: highly adaptable phytobacterial virulence factors and readily engineered DNA-targeting proteins. *Trends Cell Biol.* 23, 390–398
14. Wright, D.A. *et al.* (2014) TALEN-mediated genome editing: prospects and perspectives. *Biochem. J.* 462, 15–24
15. Belhaj, K. *et al.* (2015) Editing plant genomes with CRISPR/Cas9. *Curr. Opin. Biotechnol.* 32, 76–84
16. Luo, M. *et al.* (2016) Applications of CRISPR/Cas9 technology for targeted mutagenesis, gene replacement and stacking of genes in higher plants. *Plant Cell Rep.* 35, 1439–1450
17. Ma, X. *et al.* (2016) CRISPR/Cas9 platforms for genome editing in plants: developments and applications. *Mol. Plant.* 9, 961–974
18. Gaj, T. *et al.* (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* 31, 397–405
19. Lam, K.N. *et al.* (2011) Sequence specificity is obtained from the majority of modular C2H2 zinc-finger arrays. *Nucleic Acids Res.* 39, 4680–4690
20. Kim, H.J. *et al.* (2009) Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res.* 19, 1279–1288
21. Juillerat, A. *et al.* (2014) Comprehensive analysis of the specificity of transcription activator-like effector nucleases. *Nucleic Acids Res.* 42, 5390–5402
22. Doyle, E.L. *et al.* (2012) TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic Acids Res.* 40, W117–W122
23. Cermak, T. *et al.* (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* 39, e82
24. Upadhyay, S.K. *et al.* (2013) RNA-guided genome editing for target gene mutations in wheat. *G3 (Bethesda)* 3, 2233–2238
25. Hsu, P.D. *et al.* (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31, 827–832
26. Xie, K. *et al.* (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
27. Xie, K. and Yang, Y. (2013) RNA-guided genome editing in plants using a CRISPR-CAS system. *Mol. Plant* 6, 1975–1983
28. Wang, Y. *et al.* (2014) Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol.* 32, 947–951
29. Li, T. *et al.* (2012) High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat. Biotechnol.* 30, 390–392
30. Zhang, H. *et al.* (2016) TALEN-mediated targeted mutagenesis produces a large variety of heritable mutations in rice. *Plant Biotechnol. J.* 14, 186–194
31. Ma, X. *et al.* (2015) A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Mol. Plant* 8, 1274–1284
32. Shan, Q. *et al.* (2015) Creation of fragrant rice by targeted knockout of the OsBADH2 gene using TALEN technology. *Plant Biotechnol. J.* 13, 791–800
33. Nishizawa-Yokoi, A. *et al.* (2016) A defect in DNA Ligase4 enhances the frequency of TALEN-mediated targeted mutagenesis in rice. *Plant Physiol.* 170, 653–666
34. Baysal, C. *et al.* (2016) CRISPR/Cas9 activity in the rice *OsBEIIb* gene does not induce off-target effects in the closely related paralog *OsBEIIa*. *Mol. Breed.* 36, 83
35. Zhang, H. *et al.* (2014) The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnol. J.* 12, 797–807
36. Shan, Q. *et al.* (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 686–688
37. Xie, K. *et al.* (2015) Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3570–3575
38. Lowder, L.G. *et al.* (2015) A CRISPR/Cas9 toolbox for multiplexed plant genome editing and transcriptional regulation. *Plant Physiol.* 169, 971–985
39. Zhou, H. *et al.* (2014) Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Res.* 42, 10903–10914
40. Xu, R. *et al.* (2014) Gene targeting using the *Agrobacterium tumefaciens*-mediated CRISPR-Cas system in rice. *Rice (NY)* 7, 5
41. Mikami, M. *et al.* (2015) Parameters affecting frequency of CRISPR/Cas9 mediated targeted mutagenesis in rice. *Plant Cell Rep.* 34, 1807–1815
42. Mikami, M. *et al.* (2015) Comparison of CRISPR/Cas9 expression constructs for efficient targeted mutagenesis in rice. *Plant Mol. Biol.* 88, 561–572
43. Xu, R-F. *et al.* (2015) Generation of inheritable and "transgene clean" targeted genome-modified rice in later generations using the CRISPR/Cas9 system. *Sci. Rep.* 5, 11491
44. Li, M. *et al.* (2016) Reassessment of the four yield-related genes *Gn1a*, *DEP1*, *GS3*, and *IPA1* in rice using a CRISPR/Cas9 system. *Front. Plant Sci.* 7, 377
45. Endo, M. *et al.* (2015) Multigene knockout utilizing off-target mutations of the CRISPR/Cas9 system in rice. *Plant Cell Physiol.* 56, 41–47
46. Miao, J. *et al.* (2013) Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Res.* 23, 1233–1236
47. Wang, F. *et al.* (2016) Enhanced rice blast resistance by CRISPR/Cas9-targeted mutagenesis of the ERF transcription factor gene *OsERF922*. *PLoS ONE* 11, e0154027
48. Yang, M. *et al.* (2009) Targeted mutagenesis in the progeny of maize transgenic plants. *Plant Mol. Biol.* 70, 669–679
49. Gao, H. *et al.* (2010) Heritable targeted mutagenesis in maize using a designed endonuclease. *Plant J.* 61, 176–187
50. Djukanovic, V. *et al.* (2013) Male-sterile maize plants produced by targeted mutagenesis of the cytochrome P450-like gene (MS26) using a re-designed I-*Cre*I homing endonuclease. *Plant J.* 76, 888–899
51. Char, S.N. *et al.* (2015) Heritable site-specific mutagenesis using TALENs in maize. *Plant Biotechnol. J.* 13, 1002–1010
52. Liang, Z. *et al.* (2014) Targeted mutagenesis in *Zea mays* using TALENs and the CRISPR/Cas system. *J. Genet. Genom.* 41, 63–68
53. Svitashev, S. *et al.* (2015) Targeted mutagenesis, precise gene editing, and site-specific gene insertion in maize using Cas9 and guide RNA. *Plant Physiol.* 169, 931–945
54. Watanabe, K. *et al.* (2015) Stable gene replacement in barley by targeted double-strand break induction. *J. Exp. Bot.* 67, 1433–1445
55. Wendt, T. *et al.* (2013) TAL effector nucleases induce mutations at a pre-selected location in the genome of primary barley transformants. *Plant Mol. Biol.* 83, 279–285
56. Gurushidze, M. *et al.* (2014) True-breeding targeted gene knockout in barley using designer TALE-nuclease in haploid cells. *PLoS ONE* 9, 1–9
57. Lawrenson, T. *et al.* (2015) Induction of targeted, heritable mutations in barley and *Brassica oleracea* using RNA-guided Cas9 nuclease. *Genome Biol.* 16, 258
58. Liang, G. *et al.* (2016) Selection of highly efficient sgRNAs for CRISPR/Cas9-based plant genome editing. *Sci. Rep.* 6, 21451
59. Feng, Z. *et al.* (2013) Efficient genome editing in plants using a CRISPR/Cas system. *Cell Res.* 23, 1229–1232
60. Shukla, V.K. *et al.* (2009) Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature* 459, 437–441
61. Kim, Y. *et al.* (2013) TALENs and ZFNs are associated with different mutation signatures. *Nat. Methods* 10, 185
62. Cong, L. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823
63. Zhu, J. *et al.* (2016) Efficiency and inheritance of targeted mutagenesis in maize using CRISPR-Cas9. *J. Genet. Genom.* 43, 25–36

64. Tsai, S.Q. *et al.* (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* 33, 187–197
65. Mikami, M. *et al.* (2016) Precision targeted mutagenesis via Cas9 paired nickases in rice. *Plant Cell Physiol.* 57, 1058–1068
66. Slaymaker, I.M. *et al.* (2016) Rationally engineered Cas9 nucleases with improved specificity. *Science* 351, 84–88
67. Kleinstiver, B.P. *et al.* (2016) High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529, 490–495
68. Terada, R. *et al.* (2002) Efficient gene targeting by homologous recombination in rice. *Nat. Biotechnol.* 20, 1030–1034
69. D'Halluin, K. *et al.* (2008) Homologous recombination: a basis for targeted genome optimization in crop species such as maize. *Plant Biotechnol. J.* 6, 93–102
70. Ainley, W.M. *et al.* (2013) Trait stacking via targeted genome editing. *Plant Biotechnol. J.* 11, 1126–1134
71. Kwon, Y.I. *et al.* (2012) Overexpression of OsRecQl4 and/or OsExo1 enhances DSB-induced homologous recombination in rice. *Plant Cell Physiol.* 53, 2142–2152
72. Endo, M. *et al.* (2016) Biallelic gene targeting in rice. *Plant Physiol.* 170, 667–677
73. Sun, Y. *et al.* (2016) Engineering herbicide-resistant rice plants through CRISPR/Cas9-mediated homologous recombination of acetolactate synthase. *Mol. Plant* 9, 628–631
74. Kohli, A. *et al.* (1999) Molecular characterization of transforming plasmid rearrangements in transgenic rice reveals a recombination hotspot in the CaMV 35S promoter and confirms the predominance of microhomology mediated recombination. *Plant J.* 17, 591–601
75. Osakabe, K. *et al.* (2014) A mutated cytosine deaminase gene, codA (D314A), as an efficient negative selection marker for gene targeting in rice. *Plant Cell Physiol.* 55, 658–665
76. VanBuren, R. *et al.* (2015) Single-molecule sequencing of the desiccation-tolerant grass *Oropetium thomaeum*. *Nature* 527, 508–511
77. Brenchley, R. *et al.* (2012) Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature* 491, 705–710
78. Raats, D. *et al.* (2013) The physical map of wheat chromosome 1BS provides insights into its gene space organization and evolution. *Genome Biol.* 14, R138
79. International Barley Genome Sequencing Consortium (2012) A physical, genetic and functional sequence assembly of the barley genome. *Nature* 491, 711–716
80. Šmarda, P. *et al.* (2012) Measurements of genomic GC content in plant genomes with flow cytometry: a test for reliability. *New Phytol.* 193, 513–521
81. Shan, Q. *et al.* (2013) Rapid and efficient gene modification in rice and *Brachypodium* using TALENs. *Mol. Plant* 6, 1365–1368
82. Jiang, W. *et al.* (2013) Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Res.* 41, 1–12
83. Qiu, P. *et al.* (2004) Mutation detection using Surveyor™ nuclease. *Biotechniques* 36, 702–707
84. Vouillot, L. *et al.* (2015) Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases. *G3 (Bethesda)* 5, 407–415
85. Thomas, H.R. *et al.* (2015) High-throughput genome editing and phenotyping facilitated by high resolution melting curve analysis. *PLoS ONE* 10, e0117764
86. Brinkman, E.K. *et al.* (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* 42, e168
87. Liu, W. *et al.* (2015) DSDecode: a web-based tool for decoding of sequencing chromatograms for genotyping of targeted mutations. *Mol. Plant* 8, 1431–1433
88. Güell, M. *et al.* (2014) Genome editing assessment using CRISPR Genome Analyzer (CRISPR-GA). *Bioinformatics* 30, 2968–2970
89. Hendel, A. *et al.* (2015) Quantifying on- and off-target genome editing. *Trends Biotechnol.* 33, 132–140
90. Crosetto, N. *et al.* (2013) Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat. Methods* 10, 361–365