Title

Targeted DNA insertion in plants

Journal

Proceedings of the National Academy of Sciences of the United States of America, 118(22)

ISSN

0027-8424

Authors

Dong, Oliver Xiaoou
Ronald, Pamela C

Publication Date

2021-06-01

DOI

10.1073/pnas.2004834117

Peer reviewed

Targeted DNA insertion in plants

Oliver Xiaoou Dong[a,b,c,d,1] and Pamela C. Ronald[a,b,c,d,2]

[a]Department of Plant Pathology, University of California, Davis, CA 95616; [b]Genome Center, University of California, Davis, CA 95616; [c]Innovative Genomics Institute, Berkeley, CA 94704; and [d]Feedstocks Division, The Joint Bioenergy Institute, Emeryville, CA 94608

Edited by Dana Carroll, University of Utah, Salt Lake City, UT, and approved September 28, 2020 (received for review May 22, 2020)

Conventional methods of DNA sequence insertion into plants, using *Agrobacterium*-mediated transformation or microprojectile bombardment, result in the integration of the DNA at random sites in the genome. These plants may exhibit altered agronomic traits as a consequence of disruption or silencing of genes that serve a critical function. Also, genes of interest inserted at random sites are often not expressed at the desired level. For these reasons, targeted DNA insertion at suitable genomic sites in plants is a desirable alternative. In this paper we review approaches of targeted DNA insertion in plant genomes, discuss current technical challenges, and describe promising applications of targeted DNA insertion for crop genetic improvement.

plant genetic engineering | plant genome editing | plant genetics | targeted gene insertion | gene stacking

Providing the world's people with sufficient food and fiber while minimizing the environmental footprint of agriculture is one of the greatest challenges of our time. Genetic improvement of crop plants is an important component of enhancing the sustainability of global agricultural systems (1).

One strategy of introducing agronomically important traits into plants is through genetic engineering (Box 1), which directly manipulates the genetic makeup of the plants using molecular genetic tools (2). Compared with crop improvement methods that rely on cross-pollination, genetic engineering introduces genes encoding desirable traits directly into the plant genome. In addition, genetic engineering can be used to introduce genes from any species, expanding the diversity of agronomically useful traits that can be accessed (3).

Genetic engineering of plants often requires plant transformation (Box 1), which has been established for a wide range of plant species (4). In conventional plant transformation protocols, DNA is delivered into plant cells either via the plant-infecting soil bacterium *Agrobacterium tumefaciens* or by microprojectiles propelled by a particle gun (5). These methods result in varying copies of the DNA inserted at random locations in the host genome, which may negatively alter the plant phenotypes (6, 7). To obtain a genetically engineered plant variety with optimal phenotypes, hundreds of independent transformation events (Box 1) are typically generated (8, 9). These plants are screened for individuals carrying a single-copy insertion with high field performance (8). This pipeline can be labor-intensive and time-consuming, especially for plant species with long generation times (10).

In contrast to these conventional approaches, insertion of DNA at precharacterized genomic targets increases the chance of creating the desired traits in the resulting plants (11). During the past 30 y, various methods for targeted DNA insertion in plants have been established. Many of these methods have been improved for higher efficiencies and a broader range of genomic sites that can be targeted. In this paper, we review examples of targeted gene insertion in plants. We also discuss the technical challenges and propose strategies to address these hurdles. In the last section, we highlight the potential application of these methods in the context of agricultural production.

Methods of Targeted DNA Insertion in Plants

In this section, we provide a historical overview of methods of targeted DNA insertion in plants and provide a few examples for each method. For a more comprehensive list of the published reports of targeted DNA insertion in plants we refer readers to *SI Appendix*, Table S1.

Unaided Homologous Recombination. Pioneering research in the 1980s in mammalian cells demonstrated that exogenous DNA can be targeted to specific sites within the host genome at a low frequency through homologous recombination (HR) (12–15). These discoveries inspired the first demonstration of targeted DNA insertion in the model plant tobacco by Paszkowski et al., in 1988 (16). The researchers isolated protoplasts (Box 1) from tobacco lines carrying a partially deleted selectable marker gene and electroporated these protoplasts with DNA encoding the marker gene carrying a different deletion (16). Because the two deletions were nonoverlapping, HR between the genomic DNA and the donor DNA restored the function of the marker gene (16). The estimated efficiency of HR in this study was 0.5 to 4.2×10^{-4} (16), which was comparable to the HR frequency reported in mammalian cells around the same time (15, 17). Similar studies in tobacco and *Arabidopsis* documented HR-based targeted DNA insertion at similar efficiencies (18–21).

Researchers reported that HR-based targeted DNA insertion strategies often result in inauthentic HR events (21–24). To enrich the true HR-mediated targeted insertion events, two types of positive–negative selection systems were established. In the first system, Risseeuw et al. generated recipient tobacco plants carrying the negative selectable marker gene *codA*, which encodes a cytosine deaminase and confers lethality in the presence of the chemical compound 5- fluorocytosine (25). The researchers electroporated protoplasts derived from these recipient plants with a plasmid to induce HR in the desired manner, which would insert a kanamycin resistance marker gene and simultaneously disrupt *codA* (25). By applying kanamycin and 5-fluorocytosine at the same time, the efficiency of targeted insertion was increased to 5.7×10^{-3} (25). However, because this positive–negative selection system relied on the presence of the *codA* gene at the target site, it is not applicable to other genomic

This paper results from the NAS Colloquium of the National Academy of Sciences, "Life 2.0: The Promise and Challenge of a CRISPR Path to a Sustainable Planet," held December 10–11, 2019, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. NAS colloquia began in 1991 and have been published in PNAS since 1995. The complete program and video recordings of presentations are available on the NAS website at http://www.nasonline.org/CRISPR. The collection of colloquium papers in PNAS can be found at https://www.pnas.org/page/collection/crispr-sustainable-planet.

Author contributions: O.X.D. and P.C.R. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]Present address: State Key Laboratory of Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing, China, 210095.

[2]To whom correspondence may be addressed. Email: pcronald@ucdavis.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2004834117/-/DCSupplemental.

Published April 30, 2021.

Box 1.

Definitions of some key terms in plant genetic engineering

Genetic Engineering Genetic engineering of plants refers to the process of directly manipulating DNA with molecular genetic tools to make changes to the plant genome, often for improved agronomic traits. Examples of genetic engineering in plants include inserting DNA from a different species into the target plant genome and modifying endogenous plant DNA through genome editing.

Plant Transformation Plant transformation refers to the method of introducing DNA into the plant genome. The two conventional ways of transforming plants rely on the use of either the bacterial plant pathogen *Agrobacterium tumefaciens* or a particle gun to deliver the DNA to be inserted. These methods insert DNA at random sites in the plant genome.

Selectable Marker Genes In plant transformation, selectable marker genes are used to distinguish the cells that have uptaken the delivered DNA from the prevalent untransformed cells. Usually, the expression of a selectable marker gene confers the ability to survive a specific selection pressure. An herbicide or an antibiotic is often used during the selection process to eliminate the untransformed cells. Sometimes, a negative selectable marker gene, which confers lethality, is used to eliminate transformed cells carrying DNA inserted at unintended targets.

Plant Regeneration Plant regeneration is the process of inducing the growth and differentiation of multipotent plant cells into whole plants. Successful genetic engineering of plants often requires the regeneration of whole plants from transformed plant cells. Plant regeneration usually involves culturing plant tissue in the presence of specific plant hormones under sterile conditions.

Transformation Events Each plant regenerated from an independently transformed plant cell is considered a single transformation event. Independent transformation events vary in the location and the copy number of the DNA insert.

Protoplasts Protoplasts are plant cells with the cell wall removed. They can be transformed at high efficiency through methods of direct DNA transfer. However, regenerating plants from transformed protoplasts is often challenging.

targets. In a second system developed by Terada et al., in 2002, a dual selection plasmid carries a positive selectable marker gene (Box 1) flanked by sequence homologous to the genomic target and a diphtheria toxin gene outside the homologous region as a negative marker conferring lethality (26). The desired HR would result in the targeted insertion of only the positive selectable marker gene, whereas random integration of the delivered DNA into the host genome would result in the insertion of both the positive and the negative selectable markers (26). This approach was used to generate insertional mutants at two loci in rice at 1 to 2% efficiencies (26, 27).

Overexpressing genes known to enhance HR can also promote high-efficiency targeted DNA insertion in plants. For example, in 2005, Shaked et al. reported that constitutive expression of the yeast HR-promoting chromatin remodeling gene *Radiation sensitive 54* (*ScRAD54*) in *Arabidopsis* increased the efficiency of targeted DNA insertion by one to two orders of magnitude without altering the phenotypes (28). Notably, the insert contained a promoterless green fluorescent protein (GFP) gene without any selectable marker gene, which suggests that any nucleotide sequence can be used as the insert in this method (28).

Because of the low frequency of intrinsic HR, most of the early methods of targeted DNA insertion that relied solely on this intrinsic process were inefficient and cannot be applied broadly in plants. To overcome the inefficiency, other methods have subsequently been developed, which are described in the following sections.

Recombinase-Based Methods. Recombinases recognize specific nucleotide sequences known as recombination sites and activate the swapping of DNA (29). The relative positioning of the two recombination sites determines the outcome of the recombination (*SI Appendix*, Fig. S1*A*). When matching recombination sites are present on both the genomic target and the donor DNA carrying the nucleotide sequence to be inserted, the corresponding recombinase can catalyze the targeted insertion of the donor DNA at the genomic target (*SI Appendix*, Fig. S1 *B* and *C*). In the early 1990s, various recombinase systems were developed for site-specific gene insertion (30), such as the bacteriophage Cre-*Lox* system (31), the yeast flipase–flipase recognition target (FLP-*FRT*) system (32), and the yeast recombinase-recombination site (R-*RS*) system (33). They were exploited by plant scientists to integrate DNA fragments at defined genomic targets bearing appropriate recombination sites.

Because these recombination reactions are reversible in the presence of the recombinase, inserted DNA can be excised from the genomic target (34) (*SI Appendix*, Fig. S1*A*). This creates a challenge for the use of recombinase systems for stable on-target integration of DNA. In 1995, Albert et al. modified the Cre-*Lox* system to effectively prevent the reversal of the recombinase-catalyzed DNA insertion in tobacco plants (35). First, nucleotide sequences of *Lox* sites were altered so that the recombination reaction strongly favors one direction (35). Second, they devised strategies in which the amount of the Cre recombinase is reduced after the intended DNA insertion (35). With these improvements, a high proportion of regenerated tobacco plants carried a single-copy insertion at the designated genomic target (35). Additional examples of targeted gene insertion leading to the restoration of a marker gene using recombinase systems have been reported in tobacco (6, 36), *Arabidopsis* (37–39), soybean (40), rice (41, 42), and maize (43). These studies demonstrated that recombinase systems can be used to induce targeted gene insertion in diverse plant species.

Additional improvements of recombinase-based targeted gene insertion in plants have been made. Positive–negative selection systems have been employed in recombinase-based gene insertion methods to enrich on-target insertion events. For example, the cytokinin biosynthesis gene *isopentyl transferase* (*ipt*) or the cytosine deaminase gene *codA* have been utilized as negative selectable marker genes because they encode proteins that kill the plant tissue under the proper selective conditions (36, 38). In these experimental setups, undesired insertion of the donor plasmid at random genomic sites would introduce the negative marker gene into the plant genome and abolish the host cell (36, 38). In addition to adopting a positive–negative selection system, Nanto et al. in 2005 also placed special recombination sites on the donor DNA which promote the excision and removal of

randomly integrated donor but not the on-target DNA insert (36). With these improvements, the overall targeting efficiency reached 3% (36). Besides, improvements in plant transformation techniques have also increased the overall efficiency of recombinase-based targeted gene insertion. For example, Anand et al. in 2019 applied an optimized FLP-*FRT* system and the *Wus2-Bbm* maize transformation method (44), achieving targeted DNA insertion at a 7% efficiency in maize (43).

Recombinase-based gene insertion methods can be efficient, but they invariably depend on the availability of recipient lines carrying preintegrated recombination sites. Available genomic targets for DNA insertion are therefore limited. This limitation may be overcome by using more recent gene insertion tools (discussed below) to place recombination sites at a wider range of genomic sites as landing pads for additional targeted DNA insertion using recombinase-based methods (45).

DNA Repair-Based Methods. Genome modification at a given target can be introduced at relatively high frequencies during the repair of DNA double-stranded breaks (DSBs) (46). Cellular mechanisms to repair DSBs can be roughly classified into end joining (EJ) and HR (47, 48). The EJ pathway is further divided into nonhomologous end joining (NHEJ) and microhomology-mediated end joining (MMEJ, also known as alternative EJ) (47). During NHEJ, the two broken ends are directly rejoined, sometimes accompanied by sequence modifications at the junction ends, mostly in the form of small insertions or deletions (49). Unlike NHEJ, MMEJ involves the alignment of microhomology sequences (usually less than 16 nucleotides [nt] in length) present on both DNA ends before joining and usually results in the deletion of the nucleotide sequence between the two microhomologies (50). By contrast, the HR repair pathway is generally considered error-free and requires longer homology (usually >20 nt) on both DNA ends (47). HR is rare compared with EJ, especially in somatic plant cells (51, 52). NHEJ is widely considered as the most prevalent repair pathway in plants. NHEJ, MMEJ, and HR have all been successfully exploited for routine targeted DNA insertion (*SI Appendix*, Fig. S2) in mammalian cells (53, 54). In plants, most reported examples of targeted gene insertion through DNA repair have relied on HR. There have been only a few examples of targeted gene insertion in plants through NHEJ. We will discuss examples associated with both repair pathways in this paper. Notably, a recent study in rice suggests that the efficiencies of sequence deletion resulting from MMEJ and NHEJ are comparable, with the efficiency of MMEJ largely influenced by the availability of microhomologies near the DSB (55). This indicates that MMEJ has the potential to be exploited for targeted gene insertion in plants in the future.

Generating DSBs at defined genomic targets is crucial to efficient DNA insertion at these sites. Site-directed nucleases (SDNs) are enzymes capable of inducing DSBs at genomic targets with specific nucleotide sequences (51, 56). In this section, we highlight examples of targeted DNA insertion in plants achieved using four major SDN platforms: meganucleases, zinc finger nucleases (ZFNs), transcriptional activator-like effector nucleases (TALENs), and CRISPR-Cas. We do not include examples of SDN-mediated sequence replacement in our discussion, for which the readers are directed to two comprehensive reviews (51, 57).

Meganucleases. Meganucleases (also known as homing endonucleases) are naturally found in a large number of prokaryotic and eukaryotic organisms. They selectively cleave DNA at genomic targets with specific nucleotide sequences of 14 base pairs (bp) to 40 bp (58). The relatively high targeting specificity made them promising candidates as genome engineering tools in the 1990s.

In 1993 and 1996, Puchta et al. demonstrated that DSBs generated by the yeast meganuclease *I-SceI* can increase the frequency of HR at a specific locus by over two orders of magnitude in tobacco cells (59, 60). In the 1996 study, the researchers established an assay system in tobacco where the targeted DNA insertion by HR would activate a selectable marker gene (60). Codelivery of the donor DNA with a plasmid encoding *I-SceI*, which recognized the insertion target, boosted the HR efficiency from 10^{-5} to over 10^{-3} (60). In a subsequent study of the resulting repair junctions, *Agrobacterium*-delivered DNA encoding *I-SceI* was found to occasionally integrate at the genomic cleavage target of *I-SceI* (61). This discovery led to the application of meganucleases to insert *Agrobacterium*-delivered DNA at designated targets in the plant genome (62, 63). The use of meganucleases for targeted DNA insertion has also been demonstrated in maize (64) and barley (65) through resorting the function of a selectable marker gene.

The reliance on an existing site in the genome that matches the sequence specificity of the nuclease puts constraints on the application of meganucleases for targeted gene insertion in plants. Due to this limitation, meganucleases have largely been replaced by newer molecular tools (discussed below).

ZFNs. ZFNs are chimeric nucleases created by the fusion of a DNA-binding domain and a nonspecific DNA cleavage domain, typically derived from the endonuclease *Fok*I (66). The DNA-binding domain consists of multiple zinc finger repeats, each recognizing a distinct nucleotide triplet. By combining various zinc finger repeats, the DNA-binding domain can be programmed to recognize a specific nucleotide sequence of 9 to 18 bases (67). Because the *Fok*I domain can only cut DNA when dimerized, a pair of ZFNs that recognize sites in close proximity are used to cut DNA at the intended genomic target (68). Compared with meganucleases, ZFNs are more flexible SDNs because they can be programmed to target any genomic location.

The use of ZFNs in targeted DNA insertion in plants was first demonstrated in 2005 by Wright et al. in tobacco (69). The researchers codelivered DNA encoding a ZFN and a donor repair template into tobacco cells to repair a defective reporter gene that had previously been integrated into the tobacco genome (69). Cleavage at the defective reporter gene by the ZFN enhanced HR between the target and the repair template, resulting in the insertion of a 600-bp DNA fragment, restoring the function of the reporter gene (69). This proof-of-concept study showed that ZFNs can be used in plants to induce HR and targeted DNA insertion. Shukla et al. in 2009 used ZFNs to insert an herbicide tolerance gene within the maize metabolic gene *inositol-1,3,4,5,6-pentakisphosphate 2-kinase* (*IPK1*), disrupting its function (70). This targeted insertion conferred herbicide tolerance and reduced the accumulation of phytate, a component in seeds that can promote mineral deficiencies in humans by impairing the absorption of iron, zinc, and calcium (70).

Stacking multiple traits at a single locus is often desirable because it greatly reduces the breeding efforts. To this end, Ainley et al. in 2013 demonstrated the iterative use of ZFN-mediated targeted gene insertion through HR in maize to stack multiple marker genes at a single site (71). Building on the same system, Kumar et al. in 2015 developed a system that simultaneously exchanges selectable markers and integrates new trait genes in maize, enabling nonmarker trait genes to be stacked (72). A similar NHEJ-based gene stacking strategy by the iterative use of ZFN-mediated targeted gene insertion has also been proposed in tobacco and *Arabidopsis* (73). In another effort of stacking genes at a designated genomic target in plants, Bonawitz et al. in 2019 demonstrated the targeted insertion of a 16.2-kb DNA fragment carrying four transgenes into the soybean genome using ZFNs (74).

ZFNs are highly programmable: The amino acids of the zinc finger domain can be adjusted for a variety of genomic targets. Therefore, the application of ZFNs does not rely on the creation of recipient plant lines carrying preintroduced target sites.

https://doi.org/10.1073/pnas.2004834117

Specific zinc finger repeats have been developed for most nucleotide triplets (75). However, the modular combination of these repeats for an effective, sequence-specific ZFN requires laborious screening and optimization (75).

TALENs. Like ZFNs, TALENs are chimeric DNA-cutting enzymes resulting from the fusion of a highly modular DNA binding domain and the *Fok*I nuclease domain (76, 77). The DNA binding domain of a TALEN originates from the transcription activator-like (TAL) effectors from the bacterial plant pathogen *Xanthomonas* and consists of up to about 30 near-identical repeats (78, 79). Within each repeat, two variable amino acids dictate the recognition of a particular base on a DNA sequence (80). By joining multiple repeats, a DNA-binding domain that recognizes a specific stretch of DNA is generated (80). A pair of TALENs recognizing genomic targets in proximity can lead to the dimerization of the *Fok*I nuclease domain and induce a DSB (76, 77). TALENs can be engineered to target virtually any given DNA sequence in a relatively simple manner (77), which gives this technology additional flexibility compared with ZFNs. Notably, TALENs is the first genome editing technique used to facilitate immunotherapies by deactivating immune genes which would otherwise cause the infused immune cells to attack the patient (81). Two infants with leukemia have been successfully treated with this therapeutic approach (82), demonstrating the preciseness of TALENs in genome editing.

In 2013, Voytas and coworkers reported the in-frame insertion of a yellow fluorescent protein (*YFP*) reporter gene into an endogenous gene in tobacco protoplasts by codelivering a plasmid encoding TALENs and a repair template plasmid (83). While no fluorescence was observed when the donor plasmid was delivered alone, about 14% of the cells treated with both plasmids produced the fluorescence, indicating high-efficiency targeted gene insertion through HR (83). The same research group also used TALENs in tomato to insert a constitutive *35S* promoter upstream of the anthocyanin synthesis gene *ANT1*, which led to the accumulation of the pigment anthocyanin in regenerated tomato plants (84). In this study, DNA sequences encoding the TALENs and the repair template were placed on a viral replicon, which increases in copy number when delivered (84, 85). TALENs were also used to demonstrate targeted gene insertion in potato using the viral replicon-based delivery method, restoring the activity of a selectable marker gene or a reporter gene (86). Despite the requirement of protein engineering for every distinct target, the TALENs technology is still valued for plant genome engineering because of its programmability, efficiency, and target specificity (87).

CRISPR-Cas. The CRISPR-Cas platform originates from a prokaryotic adaptive immune system, which provides protection from invading viruses by cutting the viral nucleic acid (88). First established in 2012 as a molecular tool to cut DNA with specific nucleotide sequences (89), the CRISPR-Cas system typically consists of a Cas nuclease and a guide RNA molecule, which directs the Cas to generate DSBs at genomic targets with a defined nucleotide sequence (90). Since the report of the Cas nuclease prototype *Streptococcus pyogenes* Cas9 (SpCas9) (91), numerous naturally occurring or engineered Cas nucleases with various features have been discovered (92, 93). Target recognition by a Cas nuclease is governed by Watson–Crick base pairing between the programmable section of a guide RNA and the genomic target (89). The recognition specificity can be easily changed by modifying the variable region of the guide RNA, which makes CRISPR-Cas a highly programmable tool. The technology has been adopted in a wide range of applications (94, 95), including targeted gene insertion in many plant species (96).

In 2012, Puchta and colleagues described a strategy for targeted gene insertion in *Arabidopsis* known as *in planta* gene targeting (IPGT) (97). During IPGT, a transgenically expressed SDN simultaneously cuts the intended insertion target in the host genome and a chromosomal transgenic donor, releasing the donor DNA and causes its insertion at the intended genomic target (97). In 2014, the same group demonstrated the use of CRISPR-Cas for IPGT in *Arabidopsis* by inserting a selectable marker gene at an endogenous locus (98). In this study, the researchers first delivered the CRISPR-Cas gene and the donor DNA to plants by *Agrobacterium* as a transgenic T-DNA locus to initiate IPGT (98). After identifying the plants carrying the intended targeted insertion, the original T-DNA was removed from the genome through genetic segregation (98). With a slightly different delivery strategy, Zhao et al. in 2016 cotransformed *Arabidopsis* plants with two separate *Agrobacterium* strains, which carried the CRISPR-Cas machinery and the donor respectively, and induced IPGT causing the insertion of a GFP reporter gene at an endogenous locus (99). Although the reported insertion efficiency was low (<1%), the inserted DNA at the intended target did not contain any selectable marker gene, making the strategy applicable to the targeted insertion of virtually any DNA sequence (99). To increase the efficiency of targeted insertion of marker-free DNA, Miki et al. in 2018 reported the use of a sequential transformation approach in *Arabidopsis* to induce IPGT (100). In the first round of transformation, a plant line stably expressing *Cas9* was generated as the parental line (100). In the second round of transformation, T-DNA carrying the guide RNA and a GFP donor was delivered to the *Cas9*-expressing parental line by *Agrobacterium* to induce IPGT (100). With this method, the researchers increased the insertion frequency to between 6 and 9% without the use of chemical selection (100). The T-DNA inserts can later be removed from the plants carrying the desired DNA insert through genetic segregation.

The amount of the delivered donor DNA affects the efficiency of targeted insertion. Using an *Agrobacterium*-delivered viral replicon system (85) to enrich the donor DNA can increase the efficiency of CRISPR-Cas–mediated targeted gene insertion, as demonstrated in tomato (84, 101), potato (86), wheat (102, 103), and rice (104). Alternatively, particle bombardment, which delivers higher copies of DNA molecules to plant cells than *Agrobacterium* does, has been employed to codeliver the CRISPR-Cas machinery and the repair template into plant tissues. Targeted insertions of selectable marker genes by particle bombardment have been achieved in maize (105), soybean (106), and rice (107).

Targeted insertion of marker-free DNA has been reported in rice (108–111) and maize (112) using particle bombardment-based delivery methods. Li et al. in 2016 took advantage of the relatively more efficient NHEJ repair pathway to insert a DNA fragment into the intron of the endogenous rice gene *EPSPS*, which encodes EPSP synthase, the target of the common herbicide glyphosate (108). The inserted DNA altered the amino acid sequence of the gene product, which led to herbicide tolerance in rice (108). The mutations at the imperfect junction ends were embedded within untranslated regions of the intron and thus would not affect the protein-coding sequence (108). In another example, our team exploited the NHEJ repair mechanism to insert a DNA fragment encoding two genes involved in carotenoid biosynthesis at specific rice genomic targets, which were preevaluated for their ability to accommodate large-event mutations without altering plant performance (109). We demonstrated the targeted insertion of this carotenoid biosynthesis cassette at two independent genomic targets and obtained biofortified rice with no observable yield penalty (109). Shi et al. in 2017 applied CRISPR-Cas to insert an active promoter in front of an endogenous maize gene by HR to increase its expression (112). Increased expression of the target gene *ARGOS8* led to significantly improved grain yield under drought stress environments in the field (112). Recently, Lu et al. in 2020 demonstrated that the efficiency of NHEJ-mediated gene insertion in rice can

be enhanced by an order of magnitude when the linear donor DNA carries two specific chemical modifications at its ends, including phosphorylation and phosphorothioate linkages (111). The researchers also devised an homology-directed repair (HDR)-based sequence replacement strategy building on the method of high-frequency gene insertion (111). Notably, the methods reported in these studies do not rely on chemical selection of the DNA insert. In principle, they can be used for the site-specific insertion of any DNA sequence.

Challenges and Opportunities

Increasing On-Target Insertion Frequency. Currently, targeted DNA insertion in plants is usually inefficient, especially for large DNA fragments. Increasing the efficiency of targeted gene insertion will not only reduce the labor in screening plants but also allow the insertion of marker-free DNA fragments. To meet these goals, the frequency of targeted insertion needs to reach a practical level that allows the identification of plants with the desired insertion among a manageable population size.

Increasing the amount of donor DNA delivered can potentially promote the insertion of the donor DNA at the genomic target. Particle bombardment often delivers more DNA compared with *Agrobacterium*-mediated transformation but tends to induce more unintended sequence disruptions in the host genome (113). A large amount of donor DNA may be delivered into protoplasts with direct gene transfer methods, but regenerating plants from protoplasts is extremely challenging and remains a bottleneck for most plant species (114). Improvement in plant transformation and plant regeneration methods (Box 1) may overcome the limitations of these methods in the future. Besides, phosphorothioate linkages at the ends of a donor DNA fragment has been shown to increase the stability of the donor, which in turn contributes to increased insertion frequency (111). Also, putting the donor DNA in viral replicons (85) to increase its cellular copy number has been demonstrated as a promising strategy to achieve targeted gene insertion in a variety of plant species (84, 86, 101, 102, 104).

Biochemically tethering the donor DNA fragment to the SDN machinery leads to local donor enrichment at the insertion target, which contributes to increased insertion efficiencies in human cell lines (115, 116). Similar strategies have been used in rice for in-frame insertion of DNA encoding an HA epitope tag to label endogenous rice proteins (110, 117). In these studies, the DNA repair template was either fused with the guide RNA (117) or carried a short T-DNA border sequence, which is attracted to an *Agrobacterium* VirD2 relaxase fused to the Cas9 nuclease (110). These biochemical tethering methods have not yet been tested in other plant species.

The frequency of targeted insertion may also be increased by manipulating DNA repair pathways. DSB repair by NHEJ typically leads to the rejoining of the broken ends of the genomic target without incorporating the donor DNA. Consistently, suppressing the NHEJ repair pathway has been demonstrated to enhance DNA repair through HR in mammalian cells (118–120). Similarly, Qi et al. in 2013 reported that knocking out key NHEJ components such as *KU70* or *LIG4* in *Arabidopsis* enhanced the efficiency of HR-mediated targeted DNA insertion by three to sixteen folds (121). Similarly, a loss-of-function mutation in *LIG4* in rice was reported to shift the repair pathway from NHEJ toward MMEJ, which can potentially be employed to improve the frequency of targeted gene insertion (122). Increased HR efficiency in *Arabidopsis* has also been achieved by knocking out the nucleosome assembly gene *chromatin assembly factor 1* (*CAF-1*) (123), or by knocking out the DNA repair-related gene *Rad50*, whose homolog in yeast is involved in the cellular response to DSBs (124). HR in *Arabidopsis* can also be promoted by overexpressing the HR repair component *hypersensitive to MMS, irradiation and MMC* (*MIM*) (125) or the yeast chromatin remodeling gene *RAD54* (28). These examples demonstrate that repair pathways that favor DNA insertion can be promoted in plants by manipulating genes involved in DNA repair. These manipulations may help to achieve targeted gene insertion at higher efficiencies.

In recent years, the RNA-guided endonuclease *Lachnospiraceae bacterium* Cas12a (LbCas12a, also known as LbCpf1) has emerged as a promising tool for targeted gene insertion or sequence replacement in plants (107, 126–131). LbCas12a recognizes a T-rich protospacer-adjacent motif (PAM), which allows the nuclease to access AT-rich genomic regions where PAMs for various Cas9 nucleases are underrepresented (132). Furthermore, unlike for SpCas9, the position of the DSB induced by LbCas12a is located outside the critical region of the genomic target recognized by the guide RNA (132). Because of this feature, gene editing by LbCas12a has been hypothesized to favor large deletions, gene insertions, or gene replacements, because these changes would significantly disrupt the target sequence and prevent further cleavage by LbCas12a (132). Consistent with this hypothesis, Vu et al. in 2020 observed that the use of LbCas12a resulted in a higher frequency of targeted gene insertion compared with SpCas9 in one gene insertion experiment in tomato (126). Similarly, Wolter and Puchta in 2019 observed in an experiment in *Arabidopsis* that LbCas12a resulted in a higher IPTG efficiency compared with *Staphylococcus aureus* Cas9 (SaCas9) (129). To further test this hypothesis in plants, it is worthwhile comparing LbCas12a and Cas9 in targeted gene insertion experiments involving additional genomic targets in a wider range of plant species.

Studies in mammalian cells (133, 134) and plants (135, 136) have demonstrated that the nucleotide sequence context surrounding the DSB influences the DNA repair pathway utilized. Therefore, the choice of genomic targets for DNA insertion is crucial for high-efficiency targeted insertion in plants. However, there have been few high-throughput analyses of the DSB repair outcomes at diverse genomic sites in plants (137). Additional studies in the future may reveal potential features associated with genomic contexts that favor donor insertion during DSB repair in various plant species. This knowledge would provide guidance to target selection for increased on-target gene insertion frequencies.

Reducing Off-Target Insertions. While targeted gene insertion can occur at a reasonable efficiency in plants, off-target (ectopic) insertion of the donor DNA often arises (23, 24, 97, 109, 138). Sometimes, ectopic insertion events are mistakenly recognized as carrying the on-target DNA insert. For example, the ectopic insertion of a promoterless marker gene may unintendedly lead to its activation (65). Ectopic insertions can also generate false-positive PCR genotyping results arising from a PCR artifact known as template switching (139), giving the false impression of an on-target insertion (140). Therefore, although selection and PCR are efficient ways of identifying primary transformants, results from these assays should be verified using other methods. Southern blotting, whole-genome sequencing, or Southern-by-sequencing (141) are useful for validating the putative insert and assessing the presence of ectopic inserts. Various dual-selection systems have also proved effective in eliminating plants carrying ectopic insertions (26, 36). When both ectopic and on-target inserts exist, genetic segregation can often be used to remove the ectopic inserts.

Off-target insertions can be reduced by using SDNs with increased target specificity. For CRISPR-Cas, this can be achieved by designing guide RNAs with reduced off-target effects (142, 143) and by using Cas nucleases with enhanced specificity (144–146). Besides, using egg cell-specific or early embryo-specific promoters to drive the expression of CRISPR-Cas has

https://doi.org/10.1073/pnas.2004834117

been shown to reduce excessive nuclease activity in *Arabidopsis* (100, 147, 148).

In 2019, Liu and coworkers established a novel search-and-replace genome editing platform known as prime editing (149). In prime editing, a Cas9 nickase introduces a single-stranded break at a designated genomic target. Subsequently, a reverse transcriptase tethered to the Cas9 nickase extends the 3′ end of the nicked strand of the DNA using a programmable prime editing guide RNA as the template for reverse transcription (149). Using prime editing, the researchers introduced a wide variety of short sequence edits in human cells, including an insertion of up to 3 bp (149). Because no DSB at the editing target is incurred during prime editing, fewer off-target mutations are introduced compared with methods that involve DSB generation (149). Prime editing has been applied in genome editing in wheat (150) and rice (150–156), including the insertion of nucleotide triplets in rice cells (150, 152). As a promising genome editing platform, prime editing may be optimized in the future to achieve targeted insertion of larger DNA fragments.

Application of Targeted DNA Insertion for Crop Improvement

Genomic Safe Harbors. It has been known for decades that the expression of eukaryotic genes is influenced by the surrounding genomic context (6, 157). Appropriate spatial and temporal expression of the inserted genes depends largely on the insertion site, as the silencing of transgenes often occurs (158). In addition, the insertion of exogenous DNA may affect the expression of endogenous genes, especially when the insertion site is within a gene (159). Thus, transgene insertion at random sites through conventional transgenic approaches without proper testing often leads to unintended effects, such as decreased yield (160).

Alternatively, targeted insertion can be applied to insert trait genes into plant genomes at precharacterized sites known as genomic safe harbors, which are known not to interfere with plant performance (109). Defining genomic safe harbors for a crop plant is rewarding but can be laborious because in-depth analyses of plant performance would require multiseason field trials and the assessment of diverse phenotypes (34). Still, potential genomic safe harbors can be quickly identified by characterizing existing transgenic or mutant lines, identifying the insertion sites, measuring the expression of inserted genes, and evaluating plant performance using high-throughput phenomic approaches (161).

Stacking Multiple-Trait Genes at a Single Locus. Gene stacking in plants refers to the combination of multiple desirable trait genes, often in an elite cultivar (162). Traditionally, this is achieved through crossbreeding, where plants expressing valuable traits are cross-pollinated and individuals with combined trait genes are identified in the progeny through genetic screens. The lack of genetic linkage among individual transgenes often results in complex genetic segregation patterns. As a result, identifying plants with the desired genetic makeup often requires prolonged screening of large populations.

Alternatively, multiple genes can be positioned at a single genomic safe harbor through targeted gene insertion and crossbred into the desired cultivars as a single genetic locus. The simplified genetic segregation pattern would reduce the number of plants that need to be screened. Using targeted gene insertion for gene stacking has been demonstrated in a number of plant species (72, 163–166) and has the potential to increase the efficiency of plant breeding.

Marker-Free Insertion. In conventional plant transformation, DNA insertion into the plant genome is a rare event. Therefore, most transformation protocols rely on chemical selection to eliminate untransformed cells. Accordingly, DNA inserted through these methods must contain a selectable marker gene (Box 1), which usually remains as part of the inserted DNA in the final product and often triggers additional governmental regulation and public concern (167). In addition, from an engineering perspective, the presence of the marker in the genetic background prevents future insertion of DNA fragments containing the same marker gene. Therefore, it is desirable to generate engineered plants without selectable markers.

Because DNA insertion tends to occur at DSBs, codelivering the donor DNA with an SDN to a plant can increase the frequency of the donor DNA being inserted at the defined genomic target. Screening methods such as PCR can often be used to identify the desired insertion events among plants carrying the SDN. The process does not rely on chemical selection of the intended insert. Although a selectable marker is sometimes used to select plants carrying transgenes encoding the SDN during the delivery process, the marker gene is genetically unlinked to the target locus and can be removed from the final product through genetic segregation. This strategy of inserting marker-free DNA at designated genomic targets has been successfully applied in rice (99, 104, 108–111) and tomato (101) and may be applied to additional crop species with an established transformation protocol.

Concluding Remarks. There is an increasing need to engineer complex genetic traits into plants. For instance, in an effort to develop combined-trait corn varieties, four independent DNA inserts were combined through conventional breeding, resulting in corn plants that showed broad-spectrum insect resistance and tolerance to two herbicides at the same time (168). This massive multiyear breeding program was a joint effort of Monsanto Company and Dow AgroSciences, two of the largest agriculture companies in the United States at the time (169). In contrast, the ability to target multiple DNA elements to a single genomic site for trait stacking can simplify the inheritance pattern of the trait genes and thereby reduce the laborious breeding.

With over three decades of advancements in plant genome engineering, targeted insertion of large DNA fragments at defined genomic sites is no longer a dream. However, despite the remarkable progress in targeted insertion in plants, high-efficiency targeted insertion of large, marker-free DNA fragments and the recovery of plants with minimal off-target effects is still technically challenging. Targeted gene insertion in plants can be optimized in the future by improving DNA delivery to the plant cells, reducing the off-target effect of SDNs, and shaping the DNA repair mechanism to favor gene insertion.

New molecular tools are being developed at an accelerating rate. Beyond the conventional plant transformation methods, new delivery technologies such as carbon nanotubes (170), viral replicons (85), and de novo meristem induction (171) help overcome the hurdle of gene delivery. In addition, editing platforms based on prime editing (149, 150) or the CRISPR-Cas ribonucleoprotein (172, 173) are promising strategies with reduced off-target effects. Technological advances such as these will contribute to improved efficiency of targeted DNA insertion in plants.

Data Availability. There are no data underlying this work.

ACKNOWLEDGMENTS. We thank Michael Steinwand and Mawsheng Chern for critical reading of the manuscript. We apologize to those authors whose research could not be cited due to space limits. This work conducted by the Joint BioEnergy Institute was supported by the Office of Science, Office of Biological and Environmental Research of the US Department of Energy under contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the US Department of Energy. This work was also supported by research grants from the Innovative Genomics Institute and NIH (grant 122968) to P.C.R.

1. M. A. Steinwand, P. C. Ronald, Crop biotechnology and the future of food. *Nature Food* **1**, 273–283 (2020).
2. R. Mittler, E. Blumwald, Genetic engineering for modern agriculture: Challenges and perspectives. *Annu. Rev. Plant Biol.* **61**, 443–462 (2010).
3. S. Kumar, A. Chandra, K. C. Pandey, *Bacillus thuringiensis (Bt)* transgenic crop: An environment friendly insect-pest management strategy. *J. Environ. Biol.* **29**, 641–653 (2008).
4. F. Altpeter *et al*., Advancing crop transformation in the era of genome editing. *Plant Cell* **28**, 1510–1520 (2016).
5. A. Kohli *et al*., Transgene integration, organization and interaction in plants. *Plant Mol. Biol.* **52**, 247–258 (2003).
6. C. D. Day *et al*., Transgene integration into the same chromosome location can produce alleles that express at a predictable level, or alleles that are differentially silenced. *Genes Dev.* **14**, 2869–2880 (2000).
7. P. Meyer, Understanding and controlling transgene expression. *Trends Biotechnol.* **13**, 332–337 (1995).
8. R. H. Mumm, D. S. Walters, Quality control in the development of transgenic crop seed products. *Crop Sci.* **41**, 1381–1389 (2001).
9. Monsanto Company, "Petition for the determination of nonregulated status for increased ear biomass MON 87403 maize" (Monsanto Company, 2014).
10. R. H. Mumm, A look at product development with genetically modified crops: Examples from maize. *J. Agric. Food Chem.* **61**, 8254–8259 (2013).
11. A. Anand, T. J. Jones, Advancing agrobacterium-based crop transformation and genome modification technology for agricultural biotechnology. *Curr. Top. Microbiol. Immunol.* **418**, 489–507 (2018).
12. R. S. Kucherlapati, E. M. Eves, K. Y. Song, B. S. Morse, O. Smithies, Homologous recombination between plasmids in mammalian cells can be enhanced by treatment of input DNA. *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3153–3157 (1984).
13. K. R. Thomas, K. R. Folger, M. R. Capecchi, High frequency targeting of genes to specific sites in the mammalian genome. *Cell* **44**, 419–428 (1986).
14. T. Doetschman *et al*., Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. *Nature* **330**, 576–578 (1987).
15. K. R. Thomas, M. R. Capecchi, Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* **51**, 503–512 (1987).
16. J. Paszkowski, M. Baur, A. Bogucki, I. Potrykus, Gene targeting in plants. *EMBO J.* **7**, 4021–4026 (1988).
17. O. Smithies, R. G. Gregg, S. S. Boggs, M. A. Koralewski, R. S. Kucherlapati, Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature* **317**, 230–234 (1985).
18. R. Offringa *et al*., Extrachromosomal homologous recombination and gene targeting in plant cells after Agrobacterium mediated transformation. *EMBO J.* **9**, 3077–3084 (1990).
19. U. Halfter, P. C. Morris, L. Willmitzer, Gene targeting in *Arabidopsis thaliana*. *Mol. Gen. Genet.* **231**, 186–193 (1992).
20. M. Hrouda, J. Paszkowski, High fidelity extrachromosomal recombination and gene targeting in plants. *Mol. Gen. Genet.* **243**, 106–111 (1994).
21. E. Risseeuw, R. Offringa, M. E. Franke-van Dijk, P. J. Hooykaas, Targeted recombination in plants using *Agrobacterium* coincides with additional rearrangements at the target locus. *Plant J.* **7**, 109–119 (1995).
22. R. Offringa, M. E. Franke-van Dijk, M. J. De Groot, P. J. van den Elzen, P. J. Hooykaas, Nonreciprocal homologous recombination between Agrobacterium transferred DNA and a plant chromosomal locus. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 7346–7350 (1993).
23. M. Hanin *et al*., Gene targeting in *Arabidopsis*. *Plant J.* **28**, 671–677 (2001).
24. M. Endo, K. Osakabe, H. Ichikawa, S. Toki, Molecular characterization of true and ectopic gene targeting events at the acetolactate synthase gene in *Arabidopsis*. *Plant Cell Physiol.* **47**, 372–379 (2006).
25. E. Risseeuw, M. E. Franke-van Dijk, P. J. Hooykaas, Gene targeting and instability of *Agrobacterium* T-DNA loci in the plant genome. *Plant J.* **11**, 717–728 (1997).
26. R. Terada, H. Urawa, Y. Inagaki, K. Tsugane, S. Iida, Efficient gene targeting by homologous recombination in rice. *Nat. Biotechnol.* **20**, 1030–1034 (2002).
27. R. Terada, Y. Johzuka-Hisatomi, M. Saitoh, H. Asao, S. Iida, Gene targeting by homologous recombination as a biotechnological tool for rice functional genomics. *Plant Physiol.* **144**, 846–856 (2007).
28. H. Shaked, C. Melamed-Bessudo, A. A. Levy, High-frequency gene targeting in *Arabidopsis* plants expressing the yeast *RAD54* gene. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12265–12269 (2005).
29. S. Turan *et al*., Recombinase-mediated cassette exchange (RMCE): Traditional concepts and current challenges. *J. Mol. Biol.* **407**, 193–221 (2011).
30. D. W. Ow, The long road to recombinase-mediated plant transformation. *Plant Biotechnol. J.* **14**, 441–447 (2016).
31. P. C. Orban, D. Chui, J. D. Marth, Tissue- and site-specific DNA recombination in transgenic mice. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6861–6865 (1992).
32. T. Schlake, J. Bode, Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci. *Biochemistry* **33**, 12746–12751 (1994).
33. H. Matsuzaki, R. Nakajima, J. Nishiyama, H. Araki, Y. Oshima, Chromosome engineering in Saccharomyces cerevisiae by using a site-specific recombination system of a yeast plasmid. *J. Bacteriol.* **172**, 610–618 (1990).
34. D. W. Ow, Recombinase-directed plant transformation for the post-genomic era. *Plant Mol. Biol.* **48**, 183–200 (2002).
35. H. Albert, E. C. Dale, E. Lee, D. W. Ow, Site-specific integration of DNA into wild-type and mutant lox sites placed in the plant genome. *Plant J.* **7**, 649–659 (1995).
36. K. Nanto, K. Yamada-Watanabe, H. Ebinuma, *Agrobacterium*-mediated RMCE approach for gene replacement. *Plant Biotechnol. J.* **3**, 203–214 (2005).
37. A. C. Vergunst, L. E. Jansen, P. J. Hooykaas, Site-specific integration of *Agrobacterium* T-DNA in *Arabidopsis thaliana* mediated by Cre recombinase. *Nucleic Acids Res.* **26**, 2729–2734 (1998).
38. J. D. Louwerse *et al*., Stable recombinase-mediated cassette exchange in Arabidopsis using *Agrobacterium tumefaciens*. *Plant Physiol.* **145**, 1282–1293 (2007).
39. A. C. Vergunst, P. J. Hooykaas, Cre/lox-mediated site-specific integration of Agrobacterium T-DNA in Arabidopsis thaliana by transient expression of cre. *Plant Mol. Biol.* **38**, 393–406 (1998).
40. Z. Li *et al*., Site-specific integration of transgenes in soybean via recombinase-mediated DNA cassette exchange. *Plant Physiol.* **151**, 1087–1095 (2009).
41. V. Srivastava, D. W. Ow, Biolistic mediated site-specific integration in rice. *Mol. Breed.* **8**, 345–350 (2002).
42. S. Nandy, V. Srivastava, Site-specific gene integration in rice genome mediated by the FLP-FRT recombination system. *Plant Biotechnol. J.* **9**, 713–721 (2011).
43. A. Anand *et al*., High efficiency *Agrobacterium*-mediated site-specific gene integration in maize utilizing the FLP-*FRT* recombination system. *Plant Biotechnol. J.* **17**, 1636–1645 (2019).
44. K. Lowe *et al*., Morphogenic regulators *Baby boom* and *Wuschel* improve monocot transformation. *Plant Cell* **28**, 1998–2015 (2016).
45. H. Gao *et al*., Complex trait loci in maize enabled by CRISPR-Cas9 mediated gene insertion. *Front. Plant Sci.* **11**, 535 (2020).
46. H. Puchta, The repair of double-strand breaks in plants: Mechanisms and consequences for genome evolution. *J. Exp. Bot.* **56**, 1–14 (2005).
47. H. H. Y. Chang, N. R. Pannunzio, N. Adachi, M. R. Lieber, Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell Biol.* **18**, 495–506 (2017).
48. W. D. Wright, S. S. Shah, W. D. Heyer, Homologous recombination and the repair of DNA double-strand breaks. *J. Biol. Chem.* **293**, 10524–10535 (2018).
49. V. Gorbunova, A. A. Levy, Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. *Nucleic Acids Res.* **25**, 4650–4657 (1997).
50. A. Sfeir, L. S. Symington, Microhomology-mediated end joining: A Back-up survival mechanism or dedicated pathway? *Trends Biochem. Sci.* **40**, 701–714 (2015).
51. H. Puchta, F. Fauser, Gene targeting in plants: 25 years later. *Int. J. Dev. Biol.* **57**, 629–637 (2013).
52. Q. Que *et al*., Plant DNA repair pathways and their applications in genome engineering. *Methods Mol. Biol.* **1917**, 3–24 (2019).
53. T. Sakuma, T. Yamamoto, Magic wands of CRISPR-lots of choices for gene knock-in. *Cell Biol. Toxicol.* **33**, 501–505 (2017).
54. Y. Yamamoto, S. A. Gerbi, Making ends meet: Targeted integration of DNA fragments by genome editing. *Chromosoma* **127**, 405–420 (2018).
55. J. Tan *et al*., Efficient CRISPR/Cas9-based plant genomic fragment deletions by microhomology-mediated end joining. *Plant Biotechnol. J.*, 10.1111/pbi.13390. (2020).
56. H. Puchta, F. Fauser, Synthetic nucleases for genome engineering in plants: Prospects for a bright future. *Plant J.* **78**, 727–741 (2014).
57. Y. Sun, J. Li, L. Xia, Precise genome modification via sequence-specific nucleases-mediated gene targeting for crop improvement. *Front. Plant Sci.* **7**, 1928 (2016).
58. B. S. Chevalier, B. L. Stoddard, Homing endonucleases: Structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res.* **29**, 3757–3774 (2001).
59. H. Puchta, B. Dujon, B. Hohn, Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. *Nucleic Acids Res.* **21**, 5034–5040 (1993).
60. H. Puchta, B. Dujon, B. Hohn, Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5055–5060 (1996).
61. S. Salomon, H. Puchta, Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *EMBO J.* **17**, 6086–6095 (1998).
62. T. Tzfira, L. R. Frankman, M. Vaidya, V. Citovsky, Site-specific integration of *Agrobacterium tumefaciens* T-DNA via double-stranded intermediates. *Plant Physiol.* **133**, 1011–1023 (2003).
63. M. D. Chilton, Q. Que, Targeted integration of T-DNA into the tobacco genome at double-stranded breaks: New insights on the mechanism of T-DNA integration. *Plant Physiol.* **133**, 956–965 (2003).
64. K. D'Halluin, C. Vanderstraeten, E. Stals, M. Cornelissen, R. Ruiter, Homologous recombination: A basis for targeted genome optimization in crop species such as maize. *Plant Biotechnol. J.* **6**, 93–102 (2008).
65. K. Watanabe *et al*., Stable gene replacement in barley by targeted double-strand break induction. *J. Exp. Bot.* **67**, 1433–1445 (2016).
66. D. Carroll, Genome engineering with zinc-finger nucleases. *Genetics* **188**, 773–782 (2011).
67. D. J. Segal *et al*., Evaluation of a modular strategy for the construction of novel polydactyl zinc finger DNA-binding proteins. *Biochemistry* **42**, 2137–2148 (2003).
68. M. Bibikova *et al*., Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol. Cell. Biol.* **21**, 289–297 (2001).
69. D. A. Wright *et al*., High-frequency homologous recombination in plants mediated by zinc-finger nucleases. *Plant J.* **44**, 693–705 (2005).
70. V. K. Shukla *et al*., Precise genome modification in the crop species Zea mays using zinc-finger nucleases. *Nature* **459**, 437–441 (2009).
71. W. M. Ainley *et al*., Trait stacking via targeted genome editing. *Plant Biotechnol. J.* **11**, 1126–1134 (2013).
72. S. Kumar *et al*., A modular gene targeting system for sequential transgene stacking in plants. *J. Biotechnol.* **207**, 12–20 (2015).

73. D. M. Weinthal, R. A. Taylor, T. Tzfira, Nonhomologous end joining-mediated gene replacement in plant cells. *Plant Physiol.* **162**, 390–400 (2013).
74. N. D. Bonawitz *et al*., Zinc finger nuclease-mediated targeting of multiple transgenes to an endogenous soybean genomic locus via non-homologous end joining. *Plant Biotechnol. J.* **17**, 750–761 (2019).
75. F. D. Urnov, E. J. Rebar, M. C. Holmes, H. S. Zhang, P. D. Gregory, Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **11**, 636–646 (2010).
76. J. C. Miller *et al*., A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* **29**, 143–148 (2011).
77. F. Zhang *et al*., Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat. Biotechnol.* **29**, 149–153 (2011).
78. J. Boch *et al*., Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509–1512 (2009).
79. J. Boch, U. Bonas, Xanthomonas AvrBs3 family-type III effectors: Discovery and function. *Annu. Rev. Phytopathol.* **48**, 419–436 (2010).
80. M. J. Moscou, A. J. Bogdanove, A simple cipher governs DNA recognition by TAL effectors. *Science* **326**, 1501 (2009).
81. S. Reardon, Leukaemia success heralds wave of gene-editing therapies. *Nature* **527**, 146–147 (2015).
82. W. Qasim *et al*., Molecular remission of infant B-ALL after infusion of universal TALEN gene-edited CAR T cells. *Sci. Transl. Med.* **9**, eaaj2013 (2017).
83. Y. Zhang *et al*., Transcription activator-like effector nucleases enable efficient plant genome engineering. *Plant Physiol.* **161**, 20–27 (2013).
84. T. Čermák, N. J. Baltes, R. Čegan, Y. Zhang, D. F. Voytas, High-frequency, precise modification of the tomato genome. *Genome Biol.* **16**, 232 (2015).
85. N. J. Baltes, J. Gil-Humanes, T. Cermak, P. A. Atkins, D. F. Voytas, DNA replicons for plant genome engineering. *Plant Cell* **26**, 151–163 (2014).
86. N. M. Butler, N. J. Baltes, D. F. Voytas, D. S. Douches, Geminivirus-mediated genome editing in potato (*Solanum tuberosum* L.) using sequence-specific nucleases. *Front. Plant Sci.* **7**, 1045 (2016).
87. T. Gaj, C. A. Gersbach, C. F. Barbas 3rd, ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31**, 397–405 (2013).
88. R. Barrangou, L. A. Marraffini, CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Mol. Cell* **54**, 234–244 (2014).
89. M. Jinek *et al*., A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
90. F. Jiang, J. A. Doudna, CRISPR-Cas9 structures and mechanisms. *Annu. Rev. Biophys.* **46**, 505–529 (2017).
91. E. Deltcheva *et al*., CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602–607 (2011).
92. K. S. Makarova *et al*., Evolutionary classification of CRISPR-Cas systems: A burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **18**, 67–83 (2020).
93. W. Y. Wu, J. H. G. Lebbink, R. Kanaar, N. Geijsen, J. van der Oost, Genome editing by natural and engineered CRISPR-associated nucleases. *Nat. Chem. Biol.* **14**, 642–651 (2018).
94. H. Zhu, C. Li, C. Gao, Applications of CRISPR–Cas in agriculture and plant biotechnology. *Nat. Rev. Mol. Cell Biol.* **21**, 661–677 (2020).
95. J. A. Doudna, The promise and challenge of therapeutic genome editing. *Nature* **578**, 229–236 (2020).
96. C. Collonnier *et al*., Towards mastering CRISPR-induced gene knock-in in plants: Survey of key features and focus on the model Physcomitrella patens. *Methods* **121-122**, 103–117 (2017).
97. F. Fauser *et al*., *In planta* gene targeting. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 7535–7540 (2012).
98. S. Schiml, F. Fauser, H. Puchta, The CRISPR/Cas system can be used as nuclease for *in planta* gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. *Plant J.* **80**, 1139–1150 (2014).
99. Y. Zhao *et al*., An alternative strategy for targeted gene replacement in plants using a dual-sgRNA/Cas9 design. *Sci. Rep.* **6**, 23890 (2016).
100. D. Miki, W. Zhang, W. Zeng, Z. Feng, J. K. Zhu, CRISPR/Cas9-mediated gene targeting in Arabidopsis using sequential transformation. *Nat. Commun.* **9**, 1967 (2018).
101. T. Dahan-Meir *et al*., Efficient *in planta* gene targeting in tomato using geminiviral replicons and the CRISPR/Cas9 system. *Plant J.* **95**, 5–16 (2018).
102. J. Gil-Humanes *et al*., High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. *Plant J.* **89**, 1251–1262 (2017).
103. T. Čermák *et al*., A multipurpose toolkit to enable advanced genome engineering in plants. *Plant Cell* **29**, 1196–1217 (2017).
104. M. Wang *et al*., Gene targeting by homology-directed repair in Rice using a geminivirus-based CRISPR/Cas9 system. *Mol. Plant* **10**, 1007–1010 (2017).
105. S. Svitashev *et al*., Targeted mutagenesis, precise gene editing, and site-specific gene insertion in maize using Cas9 and guide RNA. *Plant Physiol.* **169**, 931–945 (2015).
106. Z. Li *et al*., Cas9-Guide RNA directed genome editing in soybean. *Plant Physiol.* **169**, 960–970 (2015).
107. M. B. Begemann *et al*., Precise insertion and guided editing of higher plant genomes using Cpf1 CRISPR nucleases. *Sci. Rep.* **7**, 11606 (2017).
108. J. Li *et al*., Gene replacements and insertions in rice by intron targeting using CRISPR-Cas9. *Nat. Plants* **2**, 16139 (2016).
109. O. X. Dong *et al*., Marker-free carotenoid-enriched rice generated through targeted gene insertion using CRISPR-Cas9. *Nat. Commun.* **11**, 1178 (2020).
110. Z. Ali *et al*., Fusion of the Cas9 endonuclease and the VirD2 relaxase facilitates homology-directed repair for precise genome engineering in rice. *Commun. Biol.* **3**, 44 (2020).
111. Y. Lu *et al*., Targeted, efficient sequence insertion and replacement in rice. *Nat. Biotechnol.*, 10.1038/s41587-020-0581-5 (2020).
112. J. Shi *et al*., ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. *Plant Biotechnol. J.* **15**, 207–216 (2017).
113. J. Liu *et al*., Genome-scale sequence disruption following Biolistic transformation in rice and maize. *Plant Cell* **31**, 368–383 (2019).
114. T. Eeckhaut, P. S. Lakshmanan, D. Deryckere, E. Van Bockstaele, J. Van Huylenbroeck, Progress in plant protoplast research. *Planta* **238**, 991–1003 (2013).
115. J. Carlson-Stevermer *et al*., Assembly of CRISPR ribonucleoproteins with biotinylated oligonucleotides via an RNA aptamer for precise gene editing. *Nat. Commun.* **8**, 1711 (2017).
116. E. J. Aird, K. N. Lovendahl, A. St Martin, R. S. Harris, W. R. Gordon, Increasing Cas9-mediated homology-directed repair efficiency through covalent tethering of DNA repair template. *Commun. Biol.* **1**, 54 (2018).
117. H. Butt *et al*., Efficient CRISPR/Cas9-Mediated genome editing using a chimeric single-guide RNA molecule. *Front. Plant Sci.* **8**, 1441 (2017).
118. A. J. Pierce, P. Hu, M. Han, N. Ellis, M. Jasin, Ku DNA end-binding protein modulates homologous repair of double-strand breaks in mammalian cells. *Genes Dev.* **15**, 3237–3242 (2001).
119. V. T. Chu *et al*., Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* **33**, 543–548 (2015).
120. T. Maruyama *et al*., Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* **33**, 538–542 (2015).
121. Y. Qi *et al*., Increasing frequencies of site-specific mutagenesis and gene targeting in Arabidopsis by manipulating DNA repair pathways. *Genome Res.* **23**, 547–554 (2013).
122. A. Nishizawa-Yokoi *et al*., A defect in DNA Ligase4 enhances the frequency of TALEN-mediated targeted mutagenesis in rice. *Plant Physiol.* **170**, 653–666 (2016).
123. M. Endo *et al*., Increased frequency of homologous recombination and T-DNA integration in Arabidopsis CAF-1 mutants. *EMBO J.* **25**, 5579–5590 (2006).
124. H. Gherbi *et al*., Homologous recombination *in planta* is stimulated in the absence of Rad50. *EMBO Rep.* **2**, 287–291 (2001).
125. M. Hanin, T. Mengiste, A. Bogucki, J. Paszkowski, Elevated levels of intrachromosomal homologous recombination in *Arabidopsis* overexpressing the *MIM* gene. *Plant J.* **24**, 183–189 (2000).
126. T. V. Vu *et al*., Highly efficient homology-directed repair using CRISPR/Cpf1-geminiviral replicon in tomato. *Plant Biotechnol. J.* **18**, 2133–2143 (2020).
127. S. Li, Y. Zhang, L. Xia, Y. Qi, CRISPR-Cas12a enables efficient biallelic gene targeting in rice. *Plant Biotechnol. J.* **18**, 1351–1353 (2020).
128. L. Merker, P. Schindele, T. K. Huang, F. Wolter, H. Puchta, Enhancing *in planta* gene targeting efficiencies in Arabidopsis using temperature-tolerant CRISPR/*Lb*Cas12a. *Plant Biotechnol. J.*, 10.1111/pbi.13426. (2020).
129. F. Wolter, H. Puchta, *In planta* gene targeting can be enhanced by the use of CRISPR/Cas12a. *Plant J.* **100**, 1083–1094 (2019).
130. S. Li *et al*., Synthesis-dependent repair of Cpf1-induced double strand DNA breaks enables targeted gene replacement in rice. *J. Exp. Bot.* **69**, 4715–4721 (2018).
131. S. Li *et al*., Precise gene replacement in rice by RNA transcript-templated homologous recombination. *Nat. Biotechnol.* **37**, 445–450 (2019).
132. B. Zetsche *et al*., Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771 (2015).
133. M. van Overbeek *et al*., DNA repair profiling reveals nonrandom outcomes at cas9-mediated breaks. *Mol. Cell* **63**, 633–646 (2016).
134. R. T. Leenay *et al*., Large dataset enables prediction of repair after CRISPR-Cas9 editing in primary T cells. *Nat. Biotechnol.* **37**, 1034–1037 (2019).
135. G. T. H. Vu *et al*., Endogenous sequence patterns predispose the repair modes of CRISPR/Cas9-induced DNA double-stranded breaks in *Arabidopsis thaliana*. *Plant J.* **92**, 57–67 (2017).
136. L. D. Bertier *et al*., High-resolution analysis of the efficiency, heritability, and editing outcomes of CRISPR/Cas9-Induced modifications of *NCED4* in lettuce (*Lactuca sativa*). *G3 (Bethesda)* **8**, 1513–1521 (2018).
137. H. J. Liu *et al*., High-throughput CRISPR/Cas9 mutagenesis streamlines trait gene identification in maize. *Plant Cell* **32**, 1397–1413 (2020).
138. S. de Pater, B. J. P. M. Klemann, P. J. J. Hooykaas, True gene-targeting events by CRISPR/Cas-induced DSB repair of the PPO locus with an ectopically integrated repair template. *Sci. Rep.* **8**, 3338 (2018).
139. S. J. Odelberg, R. B. Weiss, A. Hata, R. White, Template-switching during DNA synthesis by Thermus aquaticus DNA polymerase I. *Nucleic Acids Res.* **23**, 2049–2057 (1995).
140. M. Won, I. B. Dawid, PCR artifact in testing for homologous recombination in genomic editing in zebrafish. *PLoS One* **12**, e0172802 (2017).
141. G. M. Zastrow-Hayes *et al*., Southern-by-sequencing: A robust screening approach for molecular characterization of genetically modified crops. *Plant Genome* **8**, 1–15 (2015).
142. B. Minkenberg, J. Zhang, K. Xie, Y. Yang, CRISPR-PLANT v2: An online resource for highly specific guide RNA spacers based on improved off-target analysis. *Plant Biotechnol. J.* **17**, 5–8 (2019).
143. K. Xie, J. Zhang, Y. Yang, 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 (2014).
144. B. P. Kleinstiver *et al*., High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).
145. I. M. Slaymaker *et al*., Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84–88 (2016).
146. J. S. Chen *et al*., Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* **550**, 407–410 (2017).

147. Z. P. Wang *et al*., Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in *Arabidopsis* in a single generation. *Genome Biol.* **16**, 144 (2015).
148. F. Wolter, J. Klemm, H. Puchta, Efficient *in planta* gene targeting in Arabidopsis using egg cell-specific expression of the Cas9 nuclease of *Staphylococcus aureus*. *Plant J.* **94**, 735–746 (2018).
149. A. V. Anzalone *et al*., Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019).
150. Q. Lin *et al*., Prime genome editing in rice and wheat. *Nat. Biotechnol.* **38**, 582–585 (2020).
151. R. Xu *et al*., Development of plant prime-editing systems for precise genome editing. *Plant Commun.* **1**, 100043 (2020).
152. X. Tang *et al*., Plant prime editors enable precise gene editing in Rice cells. *Mol. Plant* **13**, 667–670 (2020).
153. H. Li, J. Li, J. Chen, L. Yan, L. Xia, Precise modifications of both exogenous and endogenous genes in rice by prime editing. *Mol. Plant* **13**, 671–674 (2020).
154. K. Hua, Y. Jiang, X. Tao, J. K. Zhu, Precision genome engineering in rice using prime editing system. *Plant Biotechnol. J.*, 10.1111/pbi.13395. (2020).
155. W. Xu *et al*., Versatile nucleotides substitution in plant using an improved prime editing system. *Mol. Plant* **13**, 675–678 (2020).
156. H. Butt *et al*., Engineering herbicide resistance via prime editing in rice. *Plant Biotechnol. J.*, 10.1111/pbi.13399. (2020).
157. C. Wilson, H. J. Bellen, W. J. Gehring, Position effects on eukaryotic gene expression. *Annu. Rev. Cell Biol.* **6**, 679–714 (1990).
158. K. M. J. Butaye, B. P. A. Cammue, S. L. Delauré, M. F. C. De Bolle, Approaches to minimize variation of transgene expression in plants. *Mol. Breed.* **16**, 79–91 (2005).
159. N. Bouché, D. Bouchez, *Arabidopsis* gene knockout: Phenotypes wanted. *Curr. Opin. Plant Biol.* **4**, 111–117 (2001).
160. H. Bollinedi *et al*., Molecular and functional characterization of GR2-R1 event based Backcross derived lines of golden rice in the genetic background of a mega rice variety swarna. *PLoS One* **12**, e0169600 (2017).
161. W. Yang *et al*., Crop phenomics and high-throughput phenotyping: Past decades, current challenges, and future perspectives. *Mol. Plant* **13**, 187–214 (2020).
162. Q. Que *et al*., Trait stacking in transgenic crops: Challenges and opportunities. *GM Crops* **1**, 220–229 (2010).
163. K. D'Halluin *et al*., Targeted molecular trait stacking in cotton through targeted double-strand break induction. *Plant Biotechnol. J.* **11**, 933–941 (2013).
164. L. Hou *et al*., An open-source system for *in planta* gene stacking by Bxb1 and Cre recombinases. *Mol. Plant* **7**, 1756–1765 (2014).
165. S. Kumar, A. Worden, S. Novak, R. Lee, J. F. Petolino, A trait stacking system via intragenomic homologous recombination. *Planta* **244**, 1157–1166 (2016).
166. V. Srivastava, Gene stacking in plants through the application of site-specific recombination and nuclease activity. *Methods Mol. Biol.* **1864**, 267–277 (2019).
167. Y. Y. Yau, C. N. Stewart Jr, Less is more: Strategies to remove marker genes from transgenic plants. *BMC Biotechnol.* **13**, 36 (2013).
168. D. R. Lundry, J. A. Burns, M. A. Nemeth, S. G. Riordan, Composition of grain and forage from insect-protected and herbicide-tolerant corn, MON 89034 × TC1507 × MON 88017 × DAS-59122-7 (SmartStax), is equivalent to that of conventional corn (Zea mays L.). *J. Agric. Food Chem.* **61**, 1991–1998 (2013).
169. N. P. Storer, G. D. Thompson, G. P. Head, Application of pyramided traits against Lepidoptera in insect resistance management for Bt crops. *GM Crops Food* **3**, 154–162 (2012).
170. G. S. Demirer *et al*., High aspect ratio nanomaterials enable delivery of functional genetic material without DNA integration in mature plants. *Nat. Nanotechnol.* **14**, 456–464 (2019).
171. M. F. Maher *et al*., Plant gene editing through de novo induction of meristems. *Nat. Biotechnol.* **38**, 84–89 (2020).
172. S. Svitashev, C. Schwartz, B. Lenderts, J. K. Young, A. Mark Cigan, Genome editing in maize directed by CRISPR-Cas9 ribonucleoprotein complexes. *Nat. Commun.* **7**, 13274 (2016).
173. J. W. Woo *et al*., DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat. Biotechnol.* **33**, 1162–1164 (2015).