the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2025) **121,** e70099 doi: 10.1111/tpj.70099

FOCUSED REVIEW

Beyond a few bases: methods for large DNA insertion and gene targeting in plants

Katie Vollen, Jose M. Alonso and Anna N. Stepanova*

Department of Plant Biology, North Carolina State University, Raleigh, North Carolina 27695, USA

Received 31 January 2025; revised 25 February 2025; accepted 3 March 2025.
*For correspondence (e-mail atstepan@ncsu.edu).

SUMMARY

Genome editing technologies like CRISPR/Cas have greatly accelerated the pace of both fundamental research and translational applications in agriculture. However, many plant biologists are functionally limited to creating small, targeted DNA changes or large, random DNA insertions. The ability to efficiently generate large, yet precise, DNA changes will massively accelerate crop breeding cycles, enabling researchers to more efficiently engineer crops amidst a rapidly changing agricultural landscape. This review provides an overview of existing technologies that allow plant biologists to integrate large DNA sequences within a plant host and some associated technical bottlenecks. Additionally, this review explores a selection of emerging techniques in other host systems to inspire tool development in plants.

Keywords: large DNA, genome editing, gene targeting, recombinase, integrase, BAC, transposon, T-DNA, insertion, recombination.

INTRODUCTION

As global populations continue to grow and the climate becomes increasingly unstable (Malhi et al., 2021), the need for optimization and innovation in our agricultural systems is becoming more urgent. Improved technologies for crop modification are rapidly emerging and have the potential to expedite the pace of crop breeding cycles and biological inquiry. Agricultural innovations such as selective breeding, application of industrial fertilizers and pesticides, improved irrigation and automation have resulted in large increases in average crop yield over the past few decades (Hamdan et al., 2022; Pingali, 2012). Concurrent with this initial wave of crop improvement was the development and commercialization of genetically modified (GM) crop technologies in the late 1980s through the 1990s (Abel et al., 1986; Vaeck et al., 1987). Early GM crops typically harbored single-trait transgenes to confer insect, herbicide, or viral resistance, which were randomly integrated into the genome using either *Agrobacterium*-mediated transformation or biolistic gene delivery (Hamdan et al., 2022). By 2020, GM crops comprised approximately 55% of the total harvestable cropland in the United States (USDA ERS, 2023).

In contrast to earlier GM technologies, which relied on random integration of DNA sequences into the genome, the implementation of sequence-specific nucleases (SSNs) enabled programmable genetic modifications, meaning that genetic changes could now be directed to a specific, predetermined DNA sequence. Four main types of SSNs have been adopted in plant gene editing efforts: meganucleases, zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly interspersed short palindromic repeat (CRISPR)/-Cas systems (Bhagtaney & Sundarrajan, 2023). While the structure, programmability, and applications for each SSN vary, all are capable of inducing a double-stranded break (DSB) at a specific locus. SSNs have been widely applied in plants to generate gene knockouts, facilitating the study of gene function across both model and crop species (Feng et al., 2013; Jiang et al., 2013; Zhang et al., 2010, 2013). While CRISPR/Cas9 remains the most commonly used SSN in plants, other Cas-based nucleases like Cas12 have recently been adopted in plant systems and are gaining popularity due to their expanded range of editing outcomes (Zhang et al., 2019). Recent efforts have pivoted to using SSNs to generate precise base-pair changes (via base editing or removal), epigenetic modifications, changes in gene expression, or small, precise DNA insertions by fusing a nuclease-dead version of Cas9 with

specific DNA-modifying enzymes or other effector domains (Jogam et al., 2022; Li et al., 2017; Zhu et al., 2020). Despite the successful application of SSN technologies in plants, large DNA fragment knock-ins or targeted gene replacement remain particularly challenging.

This review focuses on techniques for introducing large DNA fragments and gene targeting in plants, including (i) traditional transformation-based transgene integration methods, (ii) non-homologous end joining (NHEJ) repair-based DNA insertions/gene targeting, (iii) homology-directed repair-mediated gene insertions/gene targeting, (iv) transposon-based CRISPR/Cas-guided DNA insertions, and (v) programmable recombinase/integrase-based editing.

TRANSGENE DELIVERY SYSTEMS

The first technical hurdle that comes up when making large DNA changes or introducing transgenes into the genome is the method of DNA delivery into plant cells. Efficient transgene delivery is required for both bringing the cargo DNA intended for genomic integration and the genome-editing machinery required for the techniques outlined in this review. While several transformation techniques have been developed in plants, each comes with distinct advantages and challenges, particularly in the context of large DNA construct delivery. We examine *Agrobacterium*-mediated transformation, biolistic particle delivery, protoplast transformation, and other approaches, focusing on their technical limitations, genome-level impacts, and optimization strategies to improve transformation efficiency.

Agrobacterium-mediated transformation

Since the first introduction of a foreign gene, the glyphosate (herbicide) resistance gene (*bar*), into tobacco (*Nicotiana tobacum*) in 1983, *Agrobacterium*-mediated transformation has become the predominant method for introducing foreign genetic material into plants (Herrera-Estrella et al., 1983). *Agrobacterium tumefaciens* (referred to hereafter as *Agrobacterium*), the causal agent of crown gall disease, is a gram-negative bacterium capable of the unidirectional transfer of a segment of plasmid DNA (known as the T-DNA) into the nuclear genome of a host plant species. *Agrobacterium* strains engineered for plant transformation typically carry two separate plasmids: a disarmed Ti plasmid, which encodes virulence (*Vir*) genes that are required for *Agrobacterium* infection and T-DNA transfer (but lacks the actual T-DNA with its tumor-inducing genes), and a custom T-DNA binary vector (again, without the tumor-inducing genes) that can be modified to contain a sequence of interest (Figure 1a). The binary vector is equipped with two origins of replication and thus is amenable to replication in both *E. coli* and *Agrobacterium*. This facilitates cloning of a sequence of interest within the right and left border boundaries of the T-DNA, which will then be transferred to the plant cell (Gelvin, 2017). *Agrobacterium* containing a T-DNA of interest can then be used to infect plant tissue through floral dip transformation or *Agrobacterium*-mediated tissue culture transformation and plant regeneration (Figure 1a) (Loyola-Vargas & Ochoa-Alejo, 2018; Zhang et al., 2006).

Agrobacterium-mediated transformation is capable of transferring very large T-DNA sequences, exceeding 150 kb, into plant genomes (Alonso & Stepanova, 2014; Hamilton, 1997). However, this process lacks precision, costing researchers valuable time and resources in screening efforts. T-DNA insertions are considered random on a chromosomal level with a slight bias at the sequence level for AT-rich regions and regions that contain T-DNA border microhomology (Francis & Spiker, 2005; Kim & Gelvin, 2007; Shilo et al., 2017). As a consequence of this randomness, T-DNA insertions often result in the disruption of genes or regulatory elements that may be important for normal plant development and/or function (O'Malley & Ecker, 2010) and lead to positional effects; for example, low expression when T-DNA is inserted in transcriptionally inactive regions of chromatin (Weising et al., 1990). Furthermore, T-DNA insertions are not clean. Instead, they are often incorporated as multiple tandem copies (full or partial), sometimes containing artifacts of the plasmid backbone (Jupe et al., 2019). At the site of T-DNA integration, researchers often see insertions, deletions, and even intra/inter-chromosomal rearrangements (Jupe et al., 2019).

With the advent of affordable whole-genome sequencing, the architecture of T-DNA insertions can now be analyzed with single-nucleotide resolution. Thus, among four randomly selected T-DNA mutant lines from Arabidopsis (*Arabidopsis thaliana*) mutant collections (e.g., SALK, SAIL, WISC), long-read sequencing revealed the incorporation of as little as a 28 bp T-DNA fragment and up to seven T-DNA insertions totaling up to 236 kb in length (Jupe et al., 2019). This analysis also revealed chromosomal translocations and epigenetic changes (Jupe et al., 2019). In another report, long-read sequencing of 14 Arabidopsis GABI-Kat T-DNA insertion lines revealed an average of about two T-DNA insertions per transgenic line, in addition to instances of chromosomal translocations and truncations (Pucker et al., 2021). While these examples focused on larger structural genomic variations caused by T-DNA insertions, smaller insertions and deletions have been noted both within T-DNA insertions and at T-DNA insertion sites (Latham et al., 2006; Magembe et al., 2023). Sequencing-based strategies have also been applied in crop species to elucidate complex T-DNA insertions with similar genomic outcomes (Gang et al., 2019; Gong et al., 2021; Skarzyńska et al., 2021). These studies reveal the unpredictability and variability associated with *Agrobacterium*-mediated transformation, forcing researchers to undergo labor-intensive

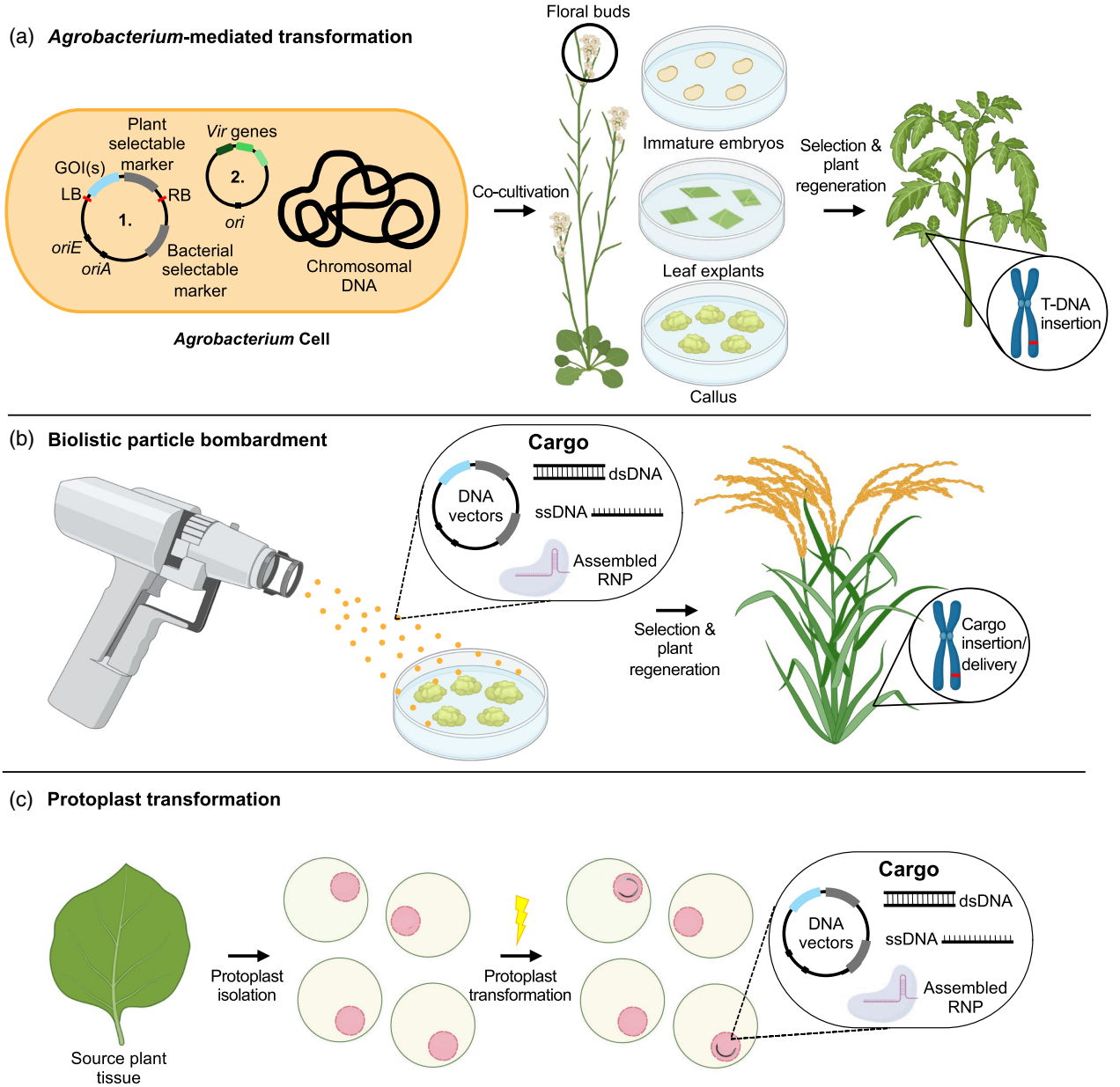

Figure 1. Transgene delivery systems.
(a) Engineered *Agrobacterium tumefaciens* strains (e.g. LBA4404, GV3101) carry a binary vector (1) and a disarmed tumor-inducing (Ti) plasmid (2). The binary vector contains the T-DNA region, flanked by left (LB) and right (RB) borders, which typically carries a gene of interest (GOI) and a selectable marker (e.g., glyphosate or kanamycin resistance) for genomic integration. The backbone of the binary vector includes dual origins of replication for plasmid maintenance in *E. coli* (oriE) and *Agrobacterium* (oriA). The disarmed Ti plasmid contains virulence (*Vir*) genes required for plant infection and T-DNA transfer. During co-cultivation, plant tissues (e.g., floral buds, leaf explants, immature embryos, or callus) are inoculated with *Agrobacterium*, facilitating T-DNA transfer via a natural infection process. Transformed tissues are placed on selective media to promote growth of only modified cells, which are regenerated into whole plants with stable T-DNA integration at random genomic sites.
(b) Biolistic particle bombardment uses a gene gun to deliver gold or tungsten particles coated with a chosen cargo (e.g., DNA vectors, double-stranded DNA, single-stranded DNA, or ribonucleoproteins) into plant tissues. Transformed cells undergo selection and regeneration to produce whole plants with random transgene integration within the genome.
(c) Isolated protoplasts (plant cells with enzymatically removed cell walls) are transformed with a cargo of choice using polyethylene glycol (PEG), electroporation, or lipofection. This approach is often used for transient expression studies, although protoplasts can sometimes be cultured and regenerated into whole plants carrying a stable transgene. Figure created in part using BioRender.

screening processes to identify lines in which a transgene is stably expressed at high levels and no major structural rearrangement is present. While such structural variations may be tolerated in basic research applications, the standards are much higher in commercial crop engineering. In industry, clean insertions are preferred, as they are

essential for regulatory approval and commercialization (Magembe et al., 2023). Despite these challenges, *Agrobacterium*-mediated transformation remains the preferred method of transformation for a wide range of dicot species and select monocots due to its ease of use and amenability to large DNA insertions.

Biolistics

Microprojectile bombardment using biolistic particle delivery emerged soon after the discovery of *Agrobacterium*-mediated plant transformation as a means of transforming plant species that are 'recalcitrant' to *Agrobacterium*-mediated transformation (Sanford, 1988). Physical means of transgene delivery, such as biolistics, are not strictly limited by genotype, thereby enabling genetic transformation of any plant species that is amenable to tissue culture and plant regeneration (Altpeter et al., 2005). Biolistic delivery involves the precipitation of nucleic acid or nucleoprotein complexes onto gold or tungsten microbeads, which are delivered to plant tissue using high velocity to penetrate the cell wall (Figure 1b) (Altpeter et al., 2005). This method has been used to deliver bacterial artificial chromosomes (BACs) with DNA cargo of up to 100 kb into potato (*Solanum tuberosum*), maize (*Zea mays*), rice (*Oryza sativa*), and wheat (*Triticum aestivum*) (Ercolano et al., 2004; Partier et al., 2017; Phan et al., 2007). Unfortunately, biolistic particle delivery often results in transgene scrambling, genomic rearrangements, and high insert copy numbers, which may lead to unstable gene expression and unpredictable consequences in native gene expression (Liu et al., 2019). For example, following biolistic delivery of a 48 kb linear fragment of lambda phage DNA into rice and maize, Liu et al., 2019 observed transformation events containing partial lambda phage DNA and multi-transgene arrays of up to 43 lambda DNA copies in rice and 51 copies in maize. In addition, these authors observed broken transgenes, large chromosomal deletions, and severe genome rearrangements. In one recovered rice transformant, sequencing revealed 1810 different lambda phage DNA fragments throughout the genome, ranging from 81 to 11 387 bp in length (Liu et al., 2019). This study highlights the severe impact biolistic transformation can have on genome structure, emphasizing the need for careful characterization of transgenic lines generated by biolistic transformation.

Other transformation methods

Protoplast transformation is another common method of plant transformation that involves enzymatic digestion of the plant cell wall and subsequent delivery of genetic material through PEG-Ca^{2+}-mediated transformation, electroporation, or lipofection (Figure 1c) (Liu et al., 2020; Reed & Bargmann, 2021; Shillito et al., 1985). Protoplast transformation is a high-efficiency method that is frequently applied in transient expression studies (Abel & Theologis, 1994). Due to the challenges associated with whole-plant regeneration from protoplasts, the isolation of stable transformants is limited to a few plant species (Reed & Bargmann, 2021). DNA vectors for protoplast transformation are typically ~5 kb or smaller in size (Burris et al., 2016; Sheen, 2001). Larger DNA vectors are typically associated with lower transformation efficiency, as they are less efficiently taken up by protoplasts (Bart et al., 2006; Ren et al., 2020). However, some protocols have been optimized to accommodate plasmids of up to 18 kb in size by tweaking tissue source/age, cell wall digestion time, incubation time, and the concentration of PEG, Ca^{2+}, and DNA (Zhang et al., 2023). Currently, protoplast transformation protocols are not well optimized to accommodate larger constructs, and the challenge of whole-plant regeneration limits the current practical applications of this technology.

Other methods of plant transformation include silicon-carbide whiskers, microinjection, pollen tube transformation, and nanomaterial technologies (Squire et al., 2023; Su et al., 2023). These transformation techniques have not yet been applied to large construct delivery and will not be covered in this review.

CRISPR/CAS9 NHEJ-MEDIATED DNA REPLACEMENT OR INSERTION INTO PLANT GENOMES

Non-homologous end joining (NHEJ) is the most common DNA double-strand break repair mechanism in most eukaryotes (Puchta, 2005; Sargent et al., 1997). Following a DSB, DNA ends are re-ligated, often resulting in small insertions or deletions (indels) at the repaired site. The error-prone nature of NHEJ has been thoroughly exploited in CRISPR/Cas9-based mutagenesis experiments, typically to generate frameshift mutations in a coding sequence (Feng et al., 2013; Jiang et al., 2013; Liang et al., 2017; Wang et al., 2014), but it can also be leveraged to incorporate DNA fragments of interest at the cut sites by providing a donor DNA molecule (Li et al., 2016).

Given the prevalence of NHEJ repair over other mechanisms, researchers sought to use NHEJ to generate large, sequence-specific insertions by targeting Cas9 modifications to the gene's intronic sequences where small indels are likely to be well tolerated (Li et al., 2016). Using particle bombardment of plasmid DNA, a pair of single guide (g) RNAs targeting two adjacent introns was delivered into rice alongside *Streptococcus pyogenes*-derived *Cas9* (*SpCas9*) and a donor DNA repair template flanked by the same pair of gRNA recognition sequences (Figure 2a). Following the Cas9-mediated DSBs in the two target sites in the genome and simultaneous DNA donor excision from the T-DNA, the genomic and donor copies were 'swapped', resulting in gene replacement via NHEJ (Figure 2a). This method was applied in rice with 2% efficiency in the 5-*ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE*

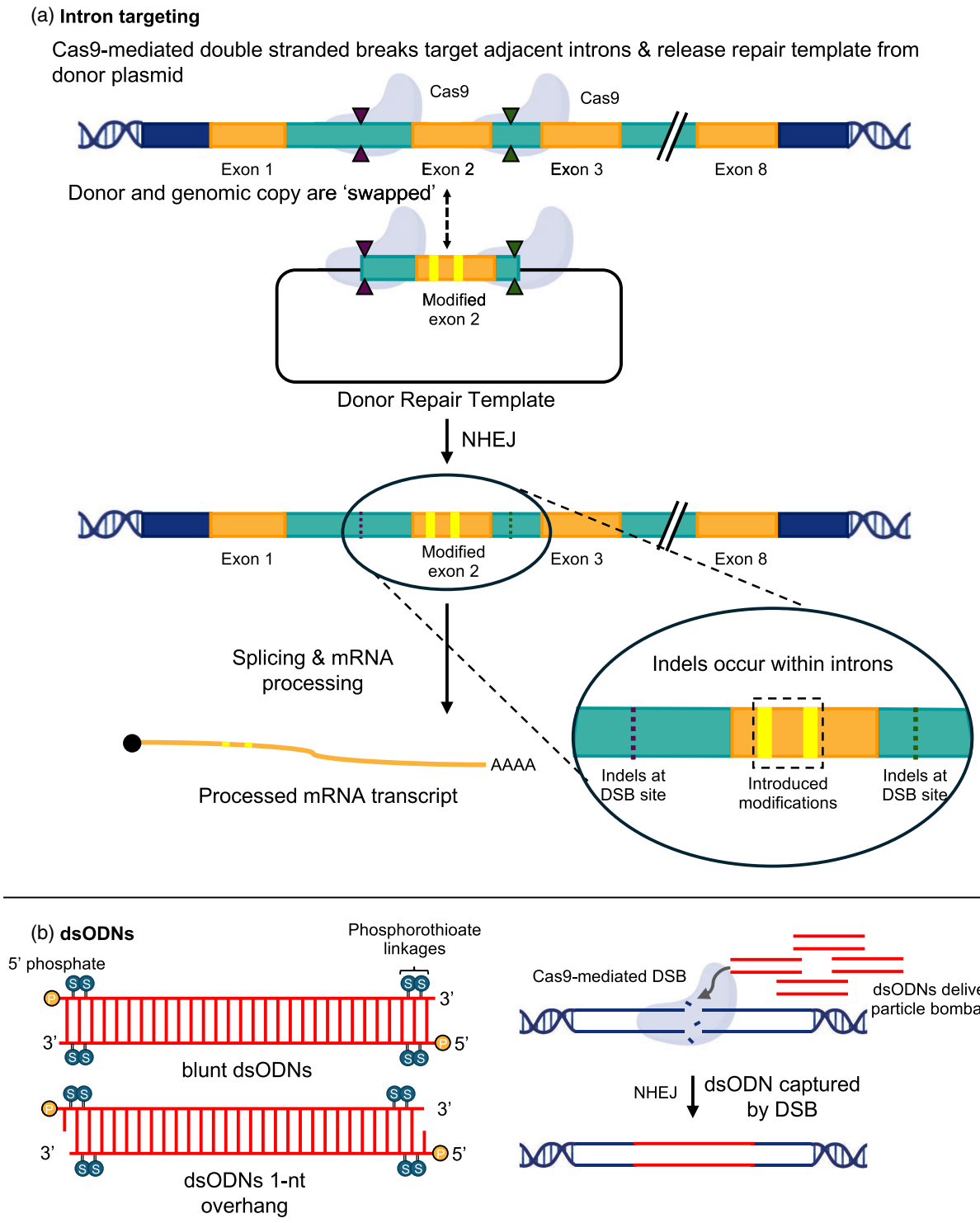

Figure 2. CRISPR/Cas9 NHEJ-mediated insertions.
(a) Intron targeting uses two gRNAs to generate double-stranded breaks (DSBs) within intronic regions on either side of the target modification site. A donor repair template is delivered in parallel, flanked by identical gRNA target sequences as the genomic target sites. Cas9-mediated cleavage results in the release of the donor repair template from the T-DNA and target site cleavage simultaneously. Following integration of the donor sequence in place of the genomic sequence, broken DNA fragments are repaired via non-homologous end joining (NHEJ), often resulting in insertions and deletions (indels) at the repaired sites. Since indels occur in intronic regions, there should be no effect on the coding sequence, and therefore, splicing and mRNA processing occur as normal.
(b) Repair template in the form of double-stranded oligodeoxynucleotides (dsODNs) with phosphorothioate modifications is delivered in excess using biolistics. The donor can either be provided as blunt-end dsODNs (top) or with 1-nt overhangs (bottom). A Cas9-mediated DSB is used to 'capture' dsODNs. DNA ends are then repaired via NHEJ.

The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd.,
The Plant Journal, (2025), **121**, e70099

(*OsEPSPS*) gene, resulting in the successful 'swapping' of a 1.6 kb fragment of exon two to introduce a double amino acid substitution that leads to glyphosate resistance. Despite the occurrence of indels at the repaired junctions, *EPSPS* was functionally spliced and translated (Li et al., 2016). However, the most frequent editing outcome was indels at the 3′ and 5′ CRISPR target sites without fragment excision or target insertion (80–83%). In instances in which cutting occurred at both gRNA sites, either fragment deletion occurred (11.3%) or the donor template was inserted in an orientation-nonspecific manner, resulting in the inversion of the genomic exon (1.8%) or inverse insertion of the donor template (1.3%) (Li et al., 2016). It is important to note, however, that even small indel mutations in an intron may impact splicing or gene regulation, requiring further screening of modified targets (Back & Walther, 2021). It remains to be seen how big of a DNA fragment can be replaced or incorporated using this approach.

Another exciting method that leverages DSB repair via NHEJ builds on previously established techniques in mammalian cells (Renaud et al., 2016). Lu et al. (2020) employed biolistics to deliver phosphorothioate-modified linear donor DNA along with a plasmid that carries CRISPR machinery to generate targeted insertions of up to 2 kb in the rice genome. An average insertion efficiency of 25% across 14 different targeted loci was achieved, with most of these (~17% of 25%) being in the desired orientation (Lu et al., 2020). The dual phosphorothioate linkages at the 5′- and 3′-ends of both DNA strands in double-stranded, 5′-phosphorylated oligodeoxynucleotides (dsODN) serve to protect the DNA from degradation in the cellular environment. Following a *SpCas9*-mediated DSB of the chromosomal DNA target, the dsODN is inserted into the genome, and the insertion junction is repaired by NHEJ (Figure 2b). This system is useful for the targeted insertions of shorter donor sequences which are compatible with commercial DNA oligo synthesis parameters (typically up to ~200 bp). Longer fragments can be generated using chemically modified oligos and PCR amplification; however, following PCR amplification, only the 5′ end will still contain a chemical modification from the annealed primer, and thus deletions at the 5′ and 3′ insertion junction would be anticipated due to DNA degradation that results from unprotected 3′ ends (Lu et al., 2020). Although this system suffers from a lack of control over the directionality of insertions and the presence of indels at insertion sites, a tweak to the dsODNs to add one-nucleotide 5′ overhangs that are complementary to the −4 position upstream of the protospacer adjacent motif (PAM) improves both issues, enabling in-frame insertions (Figure 2b). This follows recent evidence in mammals and plants that *Sp*Cas9 frequently introduces staggered cleavage, resulting in one nucleotide 5′ overhangs at the −4 position upstream of the PAM sequence (Kumar et al., 2023; Molla & Yang, 2020; Shou et al., 2018). Donor DNA molecules with one-nucleotide 5′ overhangs result in seamless targeted insertions that occur in the intended orientation with 30.1 to 60.9% frequency (Kumar et al., 2023). To demonstrate the applications of this technique, Kumar et al. applied this method to successfully tag three endogenous genes in *Setaria viridis* with the small subunit of *nanoLuciferase* (HiBiT) and the *3xFLAG* epitope, as well as engineered bacterial blight resistance in rice by inserting two distinct transcription activator-like *cis*-regulatory elements upstream of a recessive allele of the bacterial blight resistance (*R*) gene (Dixon et al., 2016). In addition, this logic could easily be extended to other nucleases that produce an overhang upon cutting, such as Cas12a or paired nickase Cas9. Overall, harnessing NHEJ repair to create targeted insertions is a promising and efficient approach that would benefit applications amenable to biolistic delivery of dsODNs. This technology may eventually be able to accommodate very large DNA insertions. As DNA synthesis technology improves, it may become possible to synthesize dsODNs beyond 2kbs or generate them through ligation of adapters containing modified bases.

It is important to distinguish between two major applications of targeted insertion strategies: instances that tolerate indels and instances that do not. In cases such as intron targeting or inserting an expression cassette into a safe-harbor region of the genome, indels are generally tolerated, making NHEJ-based approaches preferable due to their higher efficiency. In contrast, precise in-frame insertions within coding sequences require indel-free targeting and may be better suited for homology-directed repair (HDR)-based approaches due to their higher precision.

HDR-DEPENDENT PRECISION GENE TARGETING

Homologous recombination mediates the repair of damaged DNA using intact homologous chromosomal regions as a template and enables crossover events during meiosis. As discussed above, NHEJ is the dominant form of DNA repair, but it is not ideal for the generation of precise edits or large DNA insertions. In contrast, HDR is a more precise DNA repair mechanism that enables large DNA insertions or gene replacements (Figure 3a). HDR-mediated gene targeting is easily achieved in bacteria and yeast (Gardner & Jaspersen, 2014); however, it is much more challenging in higher eukaryotes due to the dominance of the NHEJ repair pathway (Sargent et al., 1997). The discovery that introducing a DSB at a target locus increases the efficiency of the HDR was the first breakthrough in the field (Puchta et al., 1993). Initial experiments were conducted by co-transforming into tobacco protoplasts the rare-cutting endonuclease, meganuclease *I-SceI*, and plasmid DNA containing two overlapping, non-functional fragments of the *uidA* (*β-glucuronidase*) gene as recombination substrates (Puchta et al., 1993). In protoplasts transformed with *I-SceI*,

functional β-glucuronidase expression, detected by histochemical staining, was 10-fold higher than in the control. Later, the utility of targeted DSB for *in planta* gene targeting (discussed below) was demonstrated in a variety of plants, including Arabidopsis, barley (*Hordeum vulgare*), and maize (Barone et al., 2020; Lawrenson et al., 2021; Schiml et al., 2014).

The introduction of programmable sequence-specific nucleases, such as ZFNs, TALENs, and the CRISPR/Cas systems, vastly expanded the practical applications of gene targeting by enabling targeted DSBs at a site of interest in the genome (Shukla et al., 2009; Townsend et al., 2009; Wright et al., 2005; Zhang et al., 2013). Still, the introduction of a DSB at a target locus alone results in a relatively low gene-targeting frequency, and existing studies often rely on the use of selectable markers to enrich for rare HDR-editing events (Butler et al., 2016; Endo et al., 2016; Kumar et al., 2016; Schiml et al., 2014; Wang et al., 2017). Several tools have since emerged to overcome some of the biological bottlenecks to favor HDR editing, as described below and as summarized in Table 1. It is important to note here that HDR efficiencies are calculated differently across studies; therefore, the reported efficiencies are often not directly comparable. Table 1 clarifies some of the methodology differences between different reports.

Ways to improve the availability of DNA repair template in HDR

Enhancing the efficiency of HDR-mediated genome editing in plants requires strategies to improve the availability of DNA repair templates at the site of a DSB. Methods, such as in planta gene targeting, leverage Cas-based nucleases to excise the DNA repair template from the genome, enabling higher rates of HDR than with a chromosomally integrated repair template alone. Further advances include tethering repair templates to SSN and utilizing viral replicons like geminiviruses to amplify repair template copy numbers, as discussed below.

Methods to mobilize DNA repair template for HDR-mediated editing

One of the first approaches to increase the efficiency of gene targeting in plants that remains relevant, known as *in planta* gene targeting (ipGT), relies on the simultaneous excision of the donor repair template from chromosomal DNA and the creation of a DSB at the target locus. The donor repair template and nuclease-encoding genes are delivered on a single T-DNA (Figure 3b). The donor repair template sequence is flanked by nuclease recognition sites, allowing the SSN to simultaneously cleave the donor repair template from the T-DNA and introduce a DSB at the target site in the genome. This coordinated cleavage mobilizes the repair template, making it more accessible at the target site. For example, following the introduction of CRISPR/Cas9, ipGT was applied in Arabidopsis using paired nickase *SpCas9* to release the DNA donor template from a T-DNA and cleave the target locus of *ALCOHOL DEHYDROGENASE1* (*AtADH1*) (Schiml et al., 2014). Subsequently, a modification of the ipGT approach involving two successive rounds of *Agrobacterium*-mediated transformation yielded a significant increase in efficiency compared to single-step transformation experiments (up to 6–9% efficiency following the second round of transformation) (Miki et al., 2018). In the Miki et al. (2018) study, the first round of transformation introduced *SpCas9* into the parental line, generating a stable line expressing only the nuclease. The second round of transformation introduced the DNA donor template and the gRNA, enabling gene editing to take place in this second round of transformation (Figure 3b). In contrast, single-step transformation experiments, in which the *SpCas9*, the gRNA, and the DNA donor template were all introduced on one T-DNA in one transformation step, resulted in no heritable gene targeting events (Miki et al., 2018).

Additional modifications to ipGT, such as the use of *Lachnospiraceae bacterium*-derived Cas12a (*LbCas12a*) or tissue-specific promoter-driven expression of *SpCas9*, further increased the efficiency of gene targeting in Arabidopsis (Miki et al., 2018; Wolter & Puchta, 2019). Given that pre-integration of *Cas9* is not always feasible or practical, an 'all-in-one' approach using heat-shock-inducible activation of *SpCas9* was used in maize (Barone et al., 2020). Following successful repair template excision from the T-DNA by Cas9 and subsequent DNA repair, an herbicide selectable marker *HIGHLY RESISTANT ACETOLACTATE SYNTHASE* (*Hra*) was constituted with its promoter element, enabling enrichment of events in which repair template excision took place (Figure 3b). This method achieved a 4.7% gene-targeting efficiency in the T0 generation (Barone et al., 2020), suggesting that enrichment for donor template excision aids in gene targeting. In another example, Li et al. employed an 'all-in-one' gene-targeting strategy using a rice codon-optimized *LbCas12a* and two distinct crRNAs to target the *ACETOLACTATE SYNTHASE* (*OsALS*) gene and cleave the donor template (Li et al., 2020). The crRNA sites were positioned approximately 484 bp apart, outside the target modification region, to increase the likelihood that at least one crRNA would successfully cleave the target site. After biolistic delivery of the 'all-in-one' vector, 1.8% of T0 events exhibited biallelic precise HDR (Li et al., 2020).

Another successful way of increasing the availability of a donor repair template at a DSB was achieved by tethering the repair template to the SSN. A 100-nucleotide single-stranded (ss) DNA repair template was delivered to rice via particle bombardment along with a vector encoding a chimeric *SpCas9-VirD2* complex (Figure 3c) (Ali et al., 2020). VirD2 is an *Agrobacterium*-derived virulence

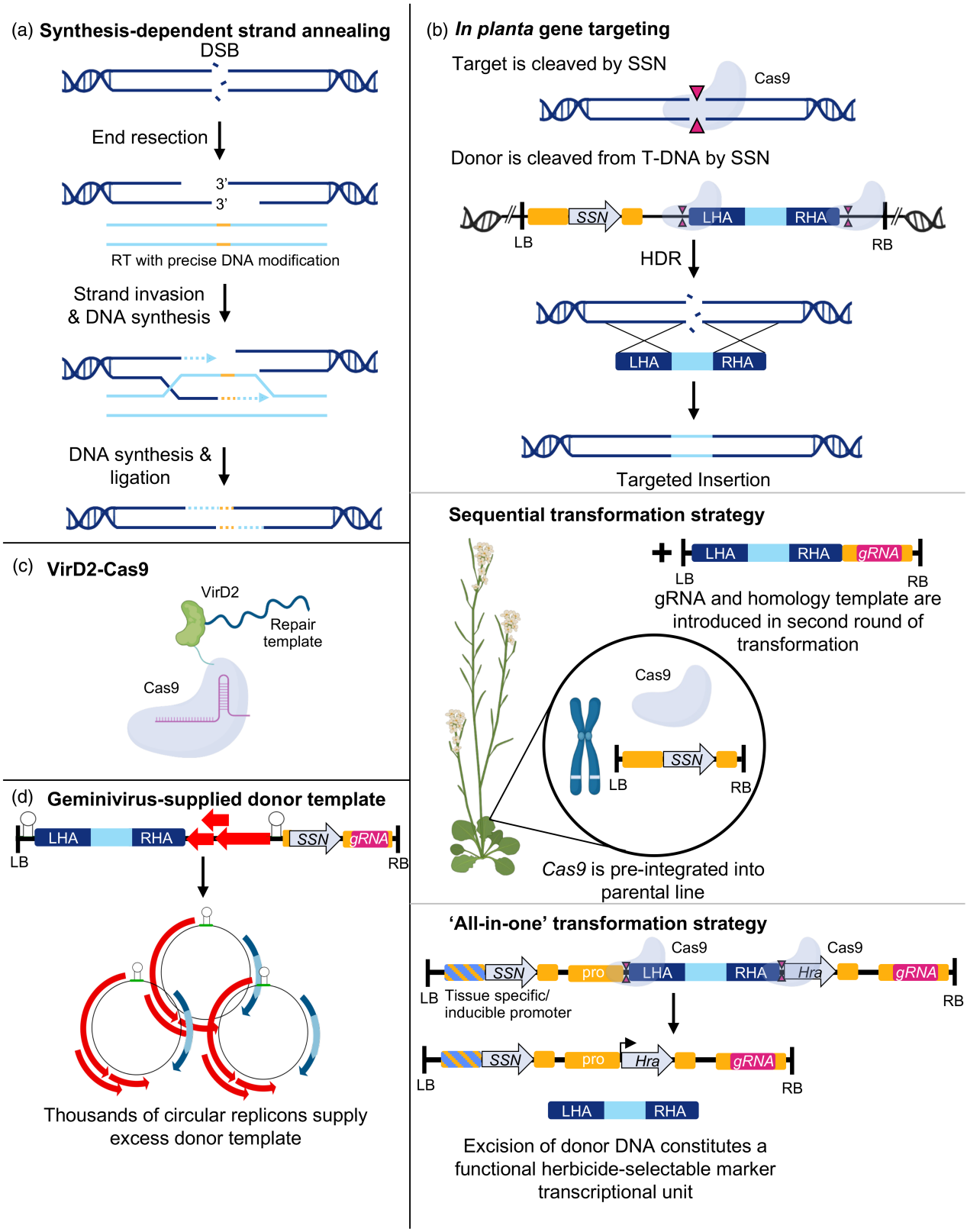

Figure 3. HDR-dependent gene targeting strategies.
(a) Overview of synthesis-dependent strand annealing (SDSA). The first step in SDSA following a double-stranded break (DSB) is end resection, creating 3′ ssDNA overhangs. Strand invasion of ssDNA overhangs initiates D-loop formation. Then DNA synthesis copies from the homologous template and fills in any remaining gaps.
(b) *In planta* gene targeting (ipGT) uses sequence-specific nucleases (SSNs) to release the donor template from the T-DNA and create a DSB at the target site. LHA, left homology arm; RHA, right homology arm. A variation of ipGT uses sequential transformation in which Cas9 is integrated into the parental plant line and gRNA and donor template are introduced in a second round of transformation. An all-in-one variation of ipGT involves just one round of transformation. The SSN is controlled by either a tissue-specific or inducible promoter (represented by striped boxes) to mediate coordinated excision of donor template and cutting at the target site. Following excision of the donor template from the T-DNA, an antibiotic or herbicide (Hra = *HIGHLY RESISTANT ACETOLACTATE SYNTHASE*) resistance transcriptional unit is constituted, enabling enrichment of events in which excision is successful.
(c) The VirD2-Cas9 system is a fusion of the *Agrobacterium* VirD2 protein (green) to Cas9. VirD2 binds a 25 nt sequence that is included in the repair template, bringing the repair template in close proximity to the Cas9 protein and target DNA break site to facilitate homology-directed repair.
(d) A geminivirus-supplied donor template strategy uses a deconstructed DNA plant virus that carries the donor template along with essential viral genes required for geminivirus-driven rolling-circle replication. This vector retains only the necessary viral sequences, excluding the coat protein and movement protein genes. Instead, the viral sequences are delivered to the plant on a T-DNA alongside a site-specific nuclease (SSN) to facilitate targeted DNA cleavage. The replication process generates thousands of circular replicons, amplifying the donor DNA and enhancing repair template availability for HDR. Figure created in part using BioRender.

protein that aids in the transfer of single-stranded T-DNA into plant cells (Dumas et al., 2001) through its ability to covalently bind to a 25-nucleotide sequence at the 5′ right border of T-DNA. By fusing *VirD2* to *Cas9*, a repair template containing the VirD2-binding sequence is drawn close to the DSB, facilitating five- to six-fold higher rates of HDR. Expression of additional *Agrobacterium* virulence genes, *VirD1* and *VirE2*, further enhanced the efficiency of the *Cas-VirD2* system, achieving 22-fold higher rates of HDR editing in rice protoplasts as compared to the delivery of *Cas9-VirD2* alone (Tang et al., 2023). Another method to tether the donor repair template to Cas proteins was demonstrated using a chimeric fusion of *LbCas12a* to the fava bean necrotic yellow virus replication (Rep) protein (a HUH endonuclease), which is capable of binding ssDNA. A four-fold increase in gene targeting was achieved in soybean compared to untethered controls (Nagy et al., 2022). It remains to be seen if this method of fusing Rep to Cas proteins can be combined with the use of geminivirus replicon proliferation to produce more DNA repair template and thus further increase the rate of HDR, as described below.

Use of geminiviruses to increase DNA repair template levels for HDR-mediated repair

Repair template availability is often considered a major bottleneck that limits the efficiency of HDR-mediated gene targeting in plants (Baltes et al., 2014). Modified plant DNA viruses are often used to increase the amount of DNA repair template accessible for gene targeting and potentially prime the plant cell into a replicative state that is more favorable for HDR repair (Figure 3d) (Baltes et al., 2014). Geminiviruses are a family of DNA viruses that can infect a wide range of crop species by reprogramming the plant host cell to favor viral replication. Pioneering work by Baltes et al. (2014) demonstrated the utility of a deconstructed geminivirus vector for the expression of SSNs and DNA repair template, achieving gene-targeting frequencies one to two orders of magnitude greater than with conventional T-DNA delivery in *Nicotiana tabacum*. The native bean yellow dwarf virus genome was modified by swapping viral coat protein and movement protein genes for DNA repair template sequence, eliminating viral cell-to-cell movement and potentially alleviating genome size constraints imposed by viral packaging and plasmodesmata size exclusion (Baltes et al., 2014; Gilbertson et al., 2003). In lieu of native viral delivery, the deconstructed vector is deployed as part of a T-DNA using *Agrobacterium*. Transformed cells accumulate thousands of copies of modified viral DNA replicons, providing an excess of repair template. Following the Baltes et al., 2014 study, many papers have been published using geminivirus vectors to deliver genome editing machinery and/or repair template in rice (Kim et al., 2022; Wang et al., 2017), cotton (*Gossypium hirsutum*) (Li et al., 2022), barley (Lawrenson et al., 2021), potato (Butler et al., 2016), grape (*Vitis vinifera*) (Olivares et al., 2021), and tomato (*Solanum lycopersicum*) (Čermák et al., 2015; Dahan-Meir et al., 2018; Vu et al., 2020). These studies, along with additional examples of geminivirus-based plant genome engineering, are summarized in Table 2, which includes details on the viral vectors used, plant species targeted, SSN employed, the viral cargo size, and editing efficiencies.

Notably, Dahan-Meir et al. (2018) achieved very high editing efficiency at the *CAROTENOID ISOMERASE* locus in tomato, with 25% of T0 plants giving rise to edited fruit in the absence of selection for HDR during the tissue-culture stage (Dahan-Meir et al., 2018). This result, however, was not replicated in other target sites, indicating that editing efficiency may be dependent on other factors such as chromatin accessibility (Filler-Hayut et al., 2021; Janssen et al., 2019). In fact, editing efficiencies achieved in different genetic loci and crop species vary greatly with this method. For example, the use of wheat dwarf virus (WDV) to modify wheat resulted in high editing efficiencies in somatic tissues, but whole plants could not be regenerated from modified cells (Gil-Humanes et al., 2017).

Table 1 Gene targeting experiments summary

Technology name	Tool delivery system	Plant model	Gene(s) targeted	Advantages	Limitations	Max size	HDR efficiency	Efficiency calculation	Citations
In planta gene targeting (ipGT)	*Agrobacterium*-mediated transformation via floral dip of a vector containing paired nickase Cas9, gRNA cassette and DNA repair template	Arabidopsis	*ALCOHOL DEHYDROGENASE1* (*ADH1*)	• Early demonstration of the use of DSBs to increase HDR efficiency	• Requires laborious screening to identify true gene targeting events • Design of paired nickases (two gRNA) constrains target site options • Antibiotic marker is needed to identify gene targeting events	Repair template was 1.8 kb long with 673 bp and 674 bp long homology arms	Two lines out of 1400 T2 seedlings	Screened 350 T2 seedlings from four independent T1 lines using PCR to identify potential gene targeting events	Fauser et al. (2012); Schiml et al. (2014)
Biolistics	Biolistic transformation of immature embryos with *Cas9*, gRNA, and repair template (ssDNA and vector)	Maize	*ACETOLACTATE SYNTHASE2* (*ALS2*), *LIGULELESS1* (*LG1*)	• Early example of gene targeting using Cas9 to generate DSB • Demonstrates potential advantages of biolistic delivery for HDR	• Uses herbicide resistance to enrich for rare editing events	Repair template was either 794 or 127 bp long containing a few point mutations; each of the homology arms were ~2.9 and ~1 kb, respectively	1.5–4.1%	Two junction PCRs were used to identify HDR events; efficiency calculated by dividing the number of events positive at both junctions by the number of calli screened	Svitashev et al. (2015)
Biolistic delivery of preassembled RNPs	Biolistic delivery of maize immature embryos with preassembled RNPs, ssDNA oligo repair template, and a helper vector containing morphogenic genes	Maize	*ALS2*	• Early demonstration of RNP delivery for gene targeting in plants	• Low efficiency, only generated heterozygous events	127 nt repair template to introduce a small point mutation (P165S)	2–2.5%	Number of callus sectors positive for gene targeting by chlorsulfuron resistance and Sanger sequencing confirmation divided by the number of bombarded embryos	Svitashev et al. (2016)
Intron targeting	Particle bombardment of rice embryogenic callus with plasmid DNA	Rice	*5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE* (*EPSPS*), *DENSE AND ERECT PANICLE1* (*DEP1*)	• Relies on NHEJ repair to increase gene targeting efficiency • Presence of indels is unlikely to affect gene expression	• Indels at gRNA sites • Undesired editing outcomes likely (native exon inversion or deletion, donor exon inversion)	1.6 kb repair template, no homology arms	2.0–2.2%	Number of T0 plants containing the replacement exon in the correct orientation per total number of T0s regenerated	Li et al. (2016)
Sequential transformation, ipGT	*Agrobacterium*-mediated transformation via floral dip; first round transformation introduces *Cas9* nuclease, second round transformation introduces gRNA and DNA donor	Arabidopsis	*REPRESSOR OF SILENCING1* (*ROS1*), *DEMETER* (*DME*), *GLABRA2* (*GL2*)	• Increased efficiency using sequential transformation strategy	• Did not detect any heritable gene targeting using an 'all-in-one' transformation strategy • Sequential strategy requires multiple rounds of transformation	The repair template was either 720 bp (*GFP*) or 1653 bp (*Luciferase*) long with homology arms ranging between ~300 and ~800 bp depending on the target gene	6–9%	Bulk T2s were grown without selection; efficiency calculated from proportion of bulk T2 lines that were positive for gene targeting; and individual plants were then screened by PCR; ~22% of individual T2s were PCR positive	Miki et al. (2018)
ipGT using egg-cell-specific expression of *SaCas9*	*Agrobacterium*-mediated transformation via floral dip of a vector containing *SaCas9*, gRNA cassette, and DNA repair template	Arabidopsis	*ALS*	• Egg-cell-specific expression of *SaCas9* has a positive impact on gene targeting frequencies	• Did not test larger DNA donors • *SaCas9* has more restrictive PAM requirements	The repair template is 12 bp long to introduce small point mutations; the homology arms are 800 and 742 bp in length	An average of ~1% of perfect gene targeting events	Efficiency is the average number of positive gene targeting lines relative to the total number of lines analyzed	Wolter et al. (2018)

(continued)

Table 1. (continued)

Technology name	Tool delivery system	Plant model	Gene(s) targeted	Advantages	Limitations	Max size	HDR efficiency	Efficiency calculation	Citations
ipGT using CRISPR/Cas12a	*Agrobacterium*-mediated transformation via floral dip of a vector containing *LbCas12a*, gRNA cassette, and DNA repair template	Arabidopsis	*ALS*	• Demonstrated that Cas12a is suitable for gene targeting	• Made use of the *ALS* resistance marker	Repair template is 41 bp long to introduce small point mutations; the homology arms are 800 and 742 bp in length	1.47% mean gene targeting rate	Efficiency is the average number of positive gene targeting lines relative to the total number of lines analyzed	Wolter and Puchta (2019)
dsODNs	Particle bombardment of rice calli with dsODNs and vector DNA encoding Cas9/gRNA	Rice	*SHOOT K+ CONCENTRATION* (*SKC1*), *DEEPER ROOTING1* (*DRO1*), *SLENDER RICE1* (*SLR1*), *SALT OVERLY SENSITIVE1* (*SOS1*)	• Efficient, simple • Relies on NHEJ to increase efficiency	• Bidirectional insertions • Indels at 5′ and 3′ junctions • Multiple tandem insertions • Donors that exceed the size limits of commercial oligo synthesis must be generated using PCR, resulting in more Indels at the insertion junctions • Restricted to biolistic delivery	2 kb repair template, higher percent seamless insertions with smaller donors (~60 bp), no homology arms	Average 17% targeted insertion in correct orientation, 6.1% precise (seamless) insertion in correct orientation	Number of T0 plants with the intended insertion divided by the total number of T0s recovered	Lu et al. (2020)
Cas9-VirD2 fusion	Particle bombardment of rice calli with plasmids containing *Cas9-VirD2* and phosphorthioate-modified ssDNA repair template	Rice	*ALS*, *CAROTENOID CLEAVAGE DIOXYGENASE7* (*CCD7*), *HISTONE DEACETYLASE* (*HDT*)	• Demonstrated heritability of modifications • Achieved editing in rice calli and whole plants	• Donor delivery is limited to particle bombardment	7 bp repair template (plus two homology arms of 50 bp) introduces small point mutations; the HA epitope repair template introduces nine amino acids as well as small point mutation; specific lengths were not shown	8–23% depending on the target locus	Allele-specific PCR was performed using one internal and one external primer; PCR product was then treated with a restriction enzyme that should only cut if the target sequence was repaired; efficiency was calculated from the frequency of digestion; Sanger sequencing and amplicon deep sequencing were used to validate some of the editing outcomes	Ali et al. (2020)
ipGT in Arabidopsis using ttLbCas12a	*Agrobacterium*-mediated transformation via floral dip of a vector containing *ttLbCas12a*, gRNA cassette, and DNA repair template	Arabidopsis	*ALS*	• A thermotolerant LbCas12a is more efficient at gene targeting than non-thermotolerant LbCas12a		Repair template introduces small point mutations; the homology arms are 637 and 905 bp in length	1.34%	Efficiency is the average number of positive gene targeting lines relative to the total number of lines analyzed	Merker et al. (2020)
LbCas12a-HUH endonuclease	Protoplast transformation	Soybean	*D5* intergenic region of chromosome 6		• Not optimized for large cargo • A large portion of editing outcomes are indels or imperfect HDR	Repair template contains 10 bp heterologous sequence; the homology arms are 30 bp long	3.3% perfect HDR	Amplicon deep sequencing to identify the percent of perfect integrations resulting from HDR; R0 plants from which at least 10% of reads contained the expected sequence modification were included in calculations	Nagy et al. (2022)

(continued)

Table 1. (continued)

Technology name	Tool delivery system	Plant model	Gene(s) targeted	Advantages	Limitations	Max size	HDR efficiency	Efficiency calculation	Citations
Gene targeting using intronized ttLbCas12a	*Agrobacterium*-mediated transformation via floral dip of a vector containing *ttLbCas12a-i*, gRNA cassette, and a repair template	Arabidopsis	*ALS*	• Demonstrates that the inclusion of introns in the *ttLbCas12a* sequence increases gene targeting efficiency	• Herbicide marker gene used	Repair template introduces small point mutations; the homology arms are 637 and 905 bp in length	2.46%	Efficiency is the average number of positive gene targeting lines relative to the total number of lines analyzed	Schindele et al. (2023)
1 bp-overhang dsODNs	PEG/Ca^{2+}–mediated protoplast transfection (Setaria) or particle bombardment (rice) of dsODNs and vector DNA encoding Cas9/gRNA	*Setaria viridis*, rice	*Male sterile 26* (*MS26*), *male sterile 26* (*MS45*), *histone H2A.W*, *Ubiquitin* (*Ubi*), *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*), *BACTERIAL BLIGHT RESISTANCE Xa23* (*BBR Xa23*)	• Efficient, simple • Relies on NHEJ to increase efficiency	• Bidirectional insertions • Indels at 5′ and 3′ junctions • Multiple tandem insertions • Donors that exceed the size limits of commercial oligo synthesis must be generated using PCR, resulting in more Indels at the insertion junctions • Restricted to biolistic delivery	66-bp FLAG epitope, no homology arms	30.1–60.9% seamless insertions in correct orientation	Number of reads with the intended insertion divided by the total number of reads with any mutation at the target site	Kumar et al. (2023)
Cas9-VirD2, VirD1, VirE2	PEG-mediated protoplast transformation and particle bombardment of embryogenic callus	Tobacco and rice	*ALS*, *phytoene desaturase* (*PDS*), *nitrate transporter 1.1B* (*NRT1.1B*)	• Shows the potential benefit of using *Agrobacterium* virulence proteins to aid gene targeting	• Did not yield homozygous edited plants • Uses visual markers or herbicide resistance to enrich for edited plants • Not optimized for large insertions, • Limited to biolistic delivery/ protoplast transformation	Repair template introduces small point mutations. Homology arms are ~100 bp in length	0.027–0.68% (rice protoplasts), 1.5% (pooled rice callus)	Amplicon deep sequencing was performed on transformed protoplasts/ transformed callus. Efficiency inferred from the percentage of reads containing precise modifications.	Tang et al. (2023)
Retrotransposon-mediated T-DNA concatenation	*Agrobacterium*-mediated transformation via floral dip of a vector containing *SaCas12a*, gRNA cassette, and DNA repair template flanked by retrotransposon-derived LIR sequences	Arabidopsis	*ALS*	• Novel mechanism to increase repair template copy number	• Used herbicide resistance to enrich for rare editing events • Retrotransposon derived sequences add ~2 kb to the repair template • High copy number insertion events may have unintended genomic and proteomic effects	Small point mutations to introduce S653N and modify the PAM site, 1.6 kb repair template with an additional ~2 kb of retrotransposon-derived sequence	4.04%	T2 seeds from individual T1 lines were bulk-analyzed using PCR primers that bind outside of the repair template. Resulting PCR products were Sanger-sequenced, and plants were determined to have undergone gene targeting if >50% of sequences had desired edits.	Dickinson et al. (2023)
PrimeRoot	PEG-mediated protoplast transformation (rice protoplasts), particle bombardment (rice calli), or *Agrobacterium*-mediated transformation (embryogenic calli)	Rice, maize	*4-HYDROXYPHENYLPYRUVATE DIOXYGENASE* (*HPPD*), *GLUTAMATE–CYSTEINE LIGASE GLUTATHIONE SYNTHETASE 1* (*GSH1*), *CELL DIVISION CYCLE 48* (*CDC48*), *ALS*, *RIBOSOMAL PROTEIN S20* (*S20*)	• Low off-target insertions, high precision • Amenable to large insertions • High efficiency	• Target site limitations inherent to dual pegRNA design restrictions • Efficiency decreases as insert size increases	Up to 11.1 kb	~1–8.4%	Efficiency calculated as either the number of integration events per 100 genomes (determined by ddPCR) or as the number of sequencing reads with seamless insertion events divided by the total number of sequencing reads	Sun et al. (2024)

(continued)

Table 1. (continued)

Technology name	Tool delivery system	Plant model	Gene(s) targeted	Advantages	Limitations	Max size	HDR efficiency	Efficiency calculation	Citations
Transposase-assisted target-site integration	*Agrobacterium*-mediated transformation	Arabidopsis, soybean	*Phytoene DESATURASE 3 (PDS3)*, *ALCOHOL DEHYDROGENASE 1 (ADH1)*, *Actin 8 (ACT8)*, *DD20* (a genomic safe harbor site in soybean)	• Very high efficiency • Amendable to large insertions	• Introduces >400 bp of *mPing* transposon in addition to the sequence of interest • Small indels and target site duplications are common • Efficiency decreases as cargo size increases • Off-target transposon insertions may occur	Up to 8.6 kb	6.5–36%	Efficiency was calculated as the percentage of T1 plants that carried the targeted insertion	Liu et al. (2024)

DSB, double-stranded break; dsDNA, double-stranded DNA; dsODNs, double-stranded oligodeoxynucleotides; HA, hemagglutinin; HDR, homology directed repair; NHEJ, non-homologous end joining; RNP, ribonucleoprotein; ssDNA, single-stranded DNA.

Additionally, to date, there are no reports on efficient HDR using geminivirus-supplied repair templates in Arabidopsis (Baltes et al., 2014; De Pater et al., 2018; Hahn et al., 2018). This is possibly due to a host-virus species incompatibility or an inefficiency associated with direct editing of germline cells as opposed to somatic cells in the case of tissue-culture-mediated editing and plant regeneration (Baltes et al., 2014; Hahn et al., 2018; Shan et al., 2018).

The typical viral vector cargo size used in these studies ranges from one to 8 kb (Baltes et al., 2014; Wang et al., 2017). Some studies report a decrease in reporter gene transcription as cargo size increases: for example, for WDV, cargo sizes exceeding 3 kb led to the poor expression of a *GFP* reporter cargo in rice calli (Wang et al., 2017). On the other hand, Gil-Humanes et al., 2017 demonstrated high-efficiency gene targeting with ~9 kb of cargo contained within the replicon boundaries of the same virus (Gil-Humanes et al., 2017). It remains unclear whether there are limitations to acceptable cargo size in coat protein and movement protein mutants of geminivirus. However, it is anticipated that viral replication may be hindered, or intra- or intermolecular recombination may occur among replicons as cargo sizes increase. To allow for larger repair template cargos in these disarmed viruses, SSN sequences can be moved out of the replicon with little impact on gene targeting efficiency (Baltes et al., 2014).

Other methods to increase DNA template availability for HDR

With biolistic delivery, transgene copy number depends on the amount of DNA delivered to cells (Altpeter et al., 2005). The multi-copy number nature of biolistic delivery can be leveraged to deliver excess DNA repair template, increasing the likelihood of HDR events (Svitashev et al., 2015). One study in maize achieved 4.1% successful gene targeting in the *ALS* gene (to confer chlorsulfuron herbicide resistance) by using biolistics to deliver *SpCas9*, gRNA, and donor template in a single vector. In contrast, when the same vector was delivered using *Agrobacterium*-mediated transformation, no successful gene-targeting events were isolated (Svitashev et al., 2015). As discussed above, there may be negative outcomes associated with biolistic delivery, such as transgene scrambling, transgene or native gene silencing, and genomic rearrangements that should be carefully considered when recovering HDR events generated using biolistics.

An alternative method developed to increase repair-template copy number relies on a retrotransposon-derived sequence in the T-DNA, resulting in the formation of large, complex T-DNA arrays (Dickinson et al., 2023). The inclusion of long terminal repeats from retrotransposon-derived sequences increased the number of T-DNA copies per genome 50-fold in Arabidopsis as compared to that of control constructs without retrotransposon sequences

Table 2 Gene targeting experiments using geminivirus vectors

Virus	Plant species	Nuclease	Max cargo size[a]	HDR efficiency	Efficiency calculation	Gene targeted	Details and limitations	Citations
Cabbage Leaf Curl Virus (CaLCUV)	Arabidopsis	ZFN	600 bp	N/A	From 23 experiments, one HDR positive plant was detected via PCR	A transgenic broken *β-glucuronidase* (*GUS*): *neomycin phosphotransferase II* (*nptII*) reporter was targeted	• HDR was not efficient in Arabidopsis using particle bombardment of whole CaLCUV virus with inducible ZFN	Baltes et al. (2014)
Bean Yellow Dwarf Virus (BeYDV)	Tobacco	ZFN, TALENs, CRISPR/*Cas9*	4.5 kb	N/A	GUS activity was used as a proxy for gene targeting efficiency	As above	• Transgenic broken *GUS:nptII* line was used to monitor HDR events in somatic tissue	Baltes et al. (2014)
BeYDV, Tomato Leaf Curl Virus (ToLCV)	Tomato	TALENs, CRISPR/*Cas9*	11.7 kb	3.65–11.66%	Efficiency is the % of cotyledons with purple spots per total number of inoculated cotyledons, normalized to the overall transformation efficiency	*ANTHOCYANIN1* (*ANT1*)	• Includes *nptII* marker to enrich for gene targeting events • Uses the overexpression phenotype of *ANT1* to visually identify gene targeting events • Heritable gene targeting events were recovered	Čermák et al. (2015)
BeYDV	Potato	TALENs, ZFN, CRISPR/*Cas9*	Not reported	12.5–41.7%	Efficiency is the % of kanamycin-tolerant events that were PCR positive for gene targeting	*Acetolactate synthase* (*ALS1*)	• Experiments were performed in a background constitutively expressing Rep • Includes *nptII* marker to enrich for gene targeting events • Only 32.2–34.5% of secondary regenerated events were positive for gene targeting	Butler et al. (2016)
Wheat Dwarf Virus (WDV)	Rice	CRISPR/*Cas9*	3 kb	7.7–19.4%	Efficiency is the % of kanamycin-tolerant events that were PCR positive for gene targeting, with gene targeting events subsequently confirmed by Sanger sequencing	*ACTIN1* (*ACT1*), *Glutathione S-transferases* (*GST1*), *GREEN FLUORESCENT PROTEIN* (*GFP*), *GUS*	• Includes *GFP* and *nptII* markers to enrich for gene targeting events • Experiments were performed in rice calli constitutively expressing *Cas9* (*Cas9+*) • Cargo >3 kb did not express reporter genes well • An all-in-one vector that includes *Cas9* was less efficient than gene targeting experiments in *Cas9+* background	Wang et al. (2017)

(continued)

Table 2. (continued)

Virus	Plant species	Nuclease	Max cargo size[a]	HDR efficiency	Efficiency calculation	Gene targeted	Details and limitations	Citations
WDV	Wheat	CRISPR/*Cas9*	9.3 kb	3.8% in wheat protoplasts, 5.74–6.4% in wheat scutella	Efficiency is the % of cells with gene targeting relative to the total transformed cells	*UBIQUITIN* (*UBI*), *MILDEW RESISTANCE LOCUS* (*MLO1*)	• Promoter-less *GFP*, *BLUE FLUORESCENT PROTEIN* (*BFP*), or *RED FLUORESCENT PROTEIN* (*RFP*) were used as visual markers to identify gene targeting events	Gil-Humanes et al. (2017)
BeYDV	Cassava	CRISPR/*Cas9*	Not reported	0% precise HR	N/A	*5-Enolpyruvylshikimate-3-phosphate synthase* (*EPSPS*)	• Editing mediated glyphosate tolerance (*EPSPS*) • Replicon design did not result in any precise HR events • Events isolated from replicon design had stunting and leaf chlorosis	Hummel et al. (2018)
BeYDV	Arabidopsis	CRISPR/*Cas9*	6.5 kb	0.0025–0.0045%	Efficiency is % of transgenic plants that carry all three T-DNAs used and survive on herbicide-(butafenacil-) containing media	*POLYPHENOL OXIDASE* (*PPO*)	• Viral replicon formation was not detected • Gene targeting efficiency was not impacted by use of a geminivirus vector • Editing mediated butafenacil tolerance (*PPO*)	De Pater et al. (2018)
BeYDV	Arabidopsis	CRISPR/*Cas9*	1.7 kb	0%	N/A	*GLABROUS1* (*GPL1*)	• No gene targeting events were isolated using a geminivirus vector	Hahn et al. (2018)
BeYDV	Tomato	CRISPR/*Cas9*	3.8 kb	25%	The percentage of T0 plants from which all fruit had undergone gene targeting, with gene targeting determined by fruit color and confirmed by PCR	*Carotenoid isomerase* (*CRISTO*), *phytoene synthase 1* (*PSY1*)	• Editing the *CRISTO* allele to visually identify gene targeting events	Dahan-Meir et al. (2018)
BeYDV	Tomato	CRISPR/*LbCas12a*	18 kb	12.8%	Efficiency is the number of purple spots per transformed cotyledon, normalized to the overall transformation efficiency	*ANT1, high-affinity K+ transporter 1, subfamily I, member 2* (*HKT1;2*)	• Multi-replicon design was implemented • An expression penalty was observed with the largest replicon sizes (18 kb)	Vu et al. (2020)
WDV	Barley	CRISPR/*Cas9*	2.7 kb	0%	N/A	*HORVU4Hr1G061310*	• No gene targeting events were isolated using WDV replicon strategy	Lawrenson et al. (2021)

(continued)

Table 2. (continued)

Virus	Plant species	Nuclease	Max cargo size[a]	HDR efficiency	Efficiency calculation	Gene targeted	Details and limitations	Citations
Beet Curly Top Virus (BCTV)	Tobacco	ZFN, CRISPR/*LbCas12a*	7.7 kb	14.3–23.5%	Efficiency is % of genomic clones in which complete gene targeting was observed from a random selection of target amplicons	*GFP, YELLOW FLUROESCENT PROTEIN (YFP)*	• Larger replicons (7.7 kb) accumulated less than smaller replicons (≤6.9 kb)	Eini et al. (2022)
BeYDV	Rice	CRISPR/*Cas9*	5.8 kb	1.32%	Percentage of transgenic calli with complete gene targeting, determined by amplicon deep sequencing	*Lycopene epsilon-cyclase* (*LcyE*)	• Relied on a visual marker	Kim et al. (2022)

CRISPR, clustered regularly interspaced short palindromic repeats; HDR, homology directed repair; PCR, polymerase chain reaction; TALENs, transcription activator-like effector nucleases, ZFN, Zinc Finger Nuclease.
[a]Max Cargo size is defined by the number of base pairs contained within the geminivirus DNA replicon.

(Dickinson et al., 2023). Retrotransposon-derived sequences can be included in the donor DNA delivery strategy to increase the amount of repair template in gene-targeting experiments. As with ipGT strategies (see above), the donor repair template is excised from an array of T-DNAs by Cas9, making multiple copies available for HDR. The inclusion of a repair template-derived sequence achieved three-fold higher gene-targeting efficiency (~4% percent true gene-targeting events at *AtALS*) than standard ipGT (Dickinson et al., 2023). However, careful analysis and outcrossing of the T-DNA construct are required to generate plants containing the desired edits but free of T-DNA insertions. Additionally, some instances of ectopic gene targeting were observed, in which parts of the genomic target sequence were copied onto the T-DNA (Dickinson et al., 2023; De Pater et al., 2018). This generated a functional gene with a repair template sequence that is randomly integrated into the genome, highlighting the importance of screening for HDR using genotyping primers that lie outside of the homology template to capture true HDR events (Dickinson et al., 2023). Besides the off-targeting concern, it is important to consider how repetitive T-DNA insertions might impact transgene expression, genome, and epigenome structure. Finally, the repair template used in this study was small (Table 1) and required the inclusion of approximately 2 kb of retrotransposon-derived sequences. The impact of these retrotransposon-derived sequences on the overall size and functionality of the repair template was not examined.

NHEJ mutants and overexpression of HDR machinery

There are both direct and indirect strategies to promote HDR-mediated gene editing. Indirect methods include suppression of the competing NHEJ pathway. Studies in plants have primarily focused on NHEJ-associated proteins like Ku70/80 and DNA ligase IV (LIG4) (Endo et al., 2016; Nishizawa-Yokoi et al., 2012; Qi et al., 2013). The Ku proteins act as a heterodimer, binding to broken DNA ends and recruiting other components of the NHEJ repair machinery, including X-RAY REPAIR CROSS-COMPLEMENTING PROTEIN4 (XRCC4) and LIG4, which help to rejoin the broken DNA strands. In Arabidopsis, gene targeting efficiency was enhanced 16-fold in a *ku70* mutant and three- to four-fold in a *lig4* mutant (Qi et al., 2013). Interestingly, in a double knockout mutant of *ku70* and the alternative NHEJ-associated *DNA polymerase Q* (*polQ*), gene-targeting efficiencies were markedly lower than in the ku70 single mutant or wild-type plants (Merker et al., 2024). Similarly, van Tol et al. (2022) reported reduced gene targeting and lack of T-DNA integration in *polQ* mutants. However, a follow-up study by Kralemann et al. (2024) found that gene-targeting efficiency was enhanced in *polQ* mutants when T-DNA was pre-integrated, suggesting an indirect effect of POLQ on gene-targeting efficiencies.

Direct promotion of HDR can also be achieved through the overexpression of HDR machinery. For example, overexpression of the yeast *DNA REPAIR AND RECOMBINATION PROTEIN54* (*RAD54*) gene in Arabidopsis enhanced gene targeting efficiency up to two-fold (Shaked et al., 2005). Furthermore, a combined strategy in which the NHEJ recombination factor, *XRCC4*, was downregulated and HDR enhancers, *C-TERMINAL–BINDING PROTEIN–INTERACTING PROTEIN* (*CtIP*) and *MEIOTIC RECOMBINATION11* (*MRE11*), were overexpressed enabled high-efficiency (48%) knockin of the bleomycin resistance gene (*BleoR*) in poplar (Movahedi et al., 2022). While efficient HDR can be achieved by either suppressing NHEJ or promoting HDR directly, it is inadvisable in most applications to tamper with the endogenous DNA repair machinery pathway to achieve precision editing. Such modifications can introduce unintended consequences, including an increased likelihood of additional mutations, growth defects, or genomic instability (Merker et al., 2024). Furthermore, in species where genetic crosses are not feasible, it may be impossible to remove these alterations post-editing, limiting their practicality for many applications. One potential approach could involve transiently knocking down NHEJ components using inducible RNA interference (RNAi). This strategy may avoid some of the complications of working in an NHEJ mutant background while still enhancing HDR efficiency (Li, Liu, et al., 2018).

Optimizing gene-targeting efficiency through nuclease selection

There is an almost constant expansion of available SSN variants for genome editing in plants (Wada et al., 2022). Herein, we will specifically highlight CRISPR-Cas systems that have been adopted in gene targeting or gene insertion experiments in plants. One such example is the application of Cas9 from *Staphylococcus aureus* (*SaCas9*) to improve the efficiency of ipGT in Arabidopsis. *SaCas9* driven by an egg-cell-specific promoter *AtEC1.1/1.2* reached gene-targeting efficiencies at the *AtALS* locus of 1–6% (Wolter et al., 2018).

Another widely used Cas enzyme is Cas12a (or Cpf1), a class 2 CRISPR family nuclease that has been widely adopted for plant genome editing applications (Dickinson et al., 2023; Merker et al., 2020; Nagy et al., 2022; Wolter & Puchta, 2019). Cas12a has expanded the range of targetable sequences to include T-rich PAM sequences that *SpCas9* is not suitable for. A codon-optimized version of *LbCas12a* that is active at lower temperatures (e.g., 22°C) more suitable for plant cultivation has been successfully applied to enhance the efficiency of ipGT in Arabidopsis, with temperature-tolerant *ttLbCas12a* boasting the highest efficiency of editing among Cas12a nucleases (1.3%) (Merker et al., 2020). Furthermore, *ttLbCas12a* was used to target the *ALS* gene in *Nicotiana tabacum*, resulting in an average gene targeting efficiency of 20% in somatic tissues and the isolation of heritable gene targeting events (Huang et al., 2021). It was postulated that Cas12a may mediate higher-efficiency gene targeting than Cas9 nucleases because cutting occurs more distally relative to the PAM sequence: if a cut is repaired through NHEJ, the PAM site likely remains intact, thereby increasing the likelihood that re-cutting can occur to give HDR a 'second chance' (Wolter & Puchta, 2019). Additionally, Cas12a cuts in a staggered manner, which results in 5′ ssDNA overhangs that may be superior for gene-targeting applications (Li, Liu, et al., 2018; Merker et al., 2020; Zhao et al., 2022).

In mammalian systems, the staggered cuts produced by *Acidaminococcus sp. Cas12a* (*AsCas12a*) have been utilized to achieve precise integrations by using a double-stranded (ds) DNA repair template with a sticky end homologous to a Cas12a overhang. This approach combines microhomology-mediated end joining (MMEJ) and HDR to facilitate targeted insertions (Zhao et al., 2022). In plants, *LbCas12a* is more widely used than *AsCas12a* due to its increased tolerance to low temperatures, which enhances editing efficiency in plant systems (Malzahn et al., 2019; Tang et al., 2017). In Arabidopsis, staggered cutting using paired nickase SpCas9 reportedly increased the efficiency of gene targeting (Čermák et al., 2017; Wolter et al., 2018). Additionally, recent evidence demonstrates that the presence of long, free 3′-ends at the target DSB can lead to enhanced gene-targeting efficiency (Schmidt et al., 2019; Schreiber et al., 2024). To generate long, free 3′-ends *in vivo*, SpCas9 or LbCas12a was fused to a herpes virus 5′ exonuclease (Schreiber et al., 2024). Gene targeting in *N. benthamiana* leaves increased up to 38-fold and up to 10-fold in Arabidopsis compared to WT Cas9 or Cas12 alone (Schreiber et al., 2024).

Finally, the use of an intronized version of *Cas9* or Cas12a can increase gene editing efficiency through intron-mediated enhancement of SSN expression (Grützner et al., 2021; Schindele et al., 2023). Grützner et al. demonstrated that the insertion of 13 introns within a maize codon-optimized version of *Cas9* (*ZmCas9i*) increased editing efficiency in Arabidopsis. While the *ZmCas9* construct without introns showed 0% editing efficiency, the intronized version achieved editing efficiencies ranging from 70 to 100% in T1 plants (Grützner et al., 2021). The insertion of introns in *ttLbCas12a-i* resulted in a more modest increase in editing, with 21–73% of primary transformants edited across seven target sites, while the intronless version's editing efficiency ranged from 5 to 57% (Schindele et al., 2023). Furthermore, Schindele et al. (2023) tested the capability of *ttLbCas12a-i* for gene targeting at the *ALS* locus in Arabidopsis, achieving 2.46% efficiency (Schindele et al., 2023).

It is expected that continuous development of SSNs will result in further improved gene-targeting efficiency, as well as increased target site flexibility through relaxed PAM site restrictions. The success of these nucleases also depends on their expression in the appropriate tissues at the right developmental stages, particularly in germline tissues, to achieve heritable gene modifications. As such, selecting the right promoter to drive SSN and gRNAs is as important as choosing the appropriate nuclease for efficient and precise genome editing (Vollen et al., 2024).

Optimizing gene targeting efficiency through promoter selection

Efficient gene targeting relies on providing an active nuclease and an accessible repair template to the right tissue at the right time. Proper selection of promoter and terminator elements to drive an SSN of interest is key to ensuring efficient nuclease expression. Constitutive promoters, such as *CaMV35S*, are often leveraged to drive ubiquitous expression of SSNs with the assumption that high transgene expression is independent of cell type. However, increasing evidence suggests that 'constitutive' promoter-driven expression varies by plant organ, developmental stage, and in response to external stimuli (Kiselev et al., 2021; Sunilkumar et al., 2002). Early gene editing experiments that made use of the *35S* promoter to drive SSN expression resulted in a high percentage of somatic and/or chimeric editing events, likely due to poor expression of *35S*-driven genes in germline tissues (Feng et al., 2013; Xing et al., 2014). For a DNA modification to be inherited in experiments in which reproductive organs are directly transformed (e.g., via floral dip), edits must occur in reproductive tissues, either the male or female gametophyte, or in early embryo development to ensure genetic changes are transmitted to offspring. Therefore, efforts have been made to identify promoters that are highly active in germline tissues for CRISPR-based experiments (Mao et al., 2016; Miki et al., 2018; Wang et al., 2015; Wolter et al., 2018; Yan et al., 2015).

In gene targeting experiments, the Arabidopsis *DOWNREGULATED IN dif1 45* (*DD45*), *CELL DIVISION CYCLE45* (*CDC45*), embryogenesis-associated *YAOZHE* (*YAO*), and the tomato pollen-specific *ANTHER-SPECIFIC PROTEIN52* (*LAT52*) gene promoters were used to drive *SpCas9* for targeted knockin of *GFP* in Arabidopsis. Among these, only *DD45* (also referred to as *EGG CELL 1.2* (*EC1.2*, *At2g21740*)), a promoter that drives expression exclusively in zygotes and early embryos, generated a high frequency of heritable, homozygous mutants by the T2 and T3 generations (Miki et al., 2018). Further comparisons of the *CLAVATA3* (*CLV3*), *YAO*, and *EC1.2/EC1.1* promoters revealed that the *EC1.2/EC1.1* enhancer-promoter fusion construct – an amalgamation of the *EC1.2* (*DD45*) enhancer and the *EC1.1* promoter – was the most effective, achieving high-efficiency gene targeting when driving *SaCas9* expression (Wolter et al., 2018). However, a subsequent study failed to reproduce high gene-targeting efficiencies using *EC1.2/EC1.1* to drive *SpCas9* (Peng et al., 2020), suggesting that other contributing factors beyond SSN expression patterns are important for determining the gene-targeting rates.

One key aspect to consider is SSN expression levels. Transcriptional and translational enhancers have been used with some success to improve *Cas9* expression and tissue specificity (Gasparis et al., 2018; Kusano et al., 2018; Peng et al., 2020). In Arabidopsis, the addition of an omega translational enhancer from tobacco mosaic virus (TMV) to the *EC1.2* (*DD45*) promoter resulted in a three-fold increase in gene targeting compared to the control without the omega sequence (Peng et al., 2020). The addition of the *35S* enhancer to the *EC1.1* promoter increased *Cas9* expression but in a non-tissue-specific manner. In contrast, the addition of the *EC1.2* (*DD45*) enhancer to the *EC1.1* promoter resulted in higher rates of gene editing (Wang et al., 2015). Interestingly, the incorporation of the omega translational enhancer into the *EC1.2en* (*DD45*)/*EC1.1* enhancer-promoter fusion construct did not yield any gene targeting events (Wang et al., 2015). Additionally, this study demonstrated the importance of terminator selection by comparing combinations of eight promoters and two terminators: the *Agrobacterium*-derived *nos* terminator and the *Pisum sativum*-derived *SMALL SUBUNIT OF RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE* (*rbcS*)-*E9* terminator. The *rbcS E9t* terminator combinations resulted in more heritable editing events overall than any *nos* terminator combinations, potentially due to increased mRNA stability resulting from strong transcriptional termination in the target tissues (Wang et al., 2015).

Even modest improvements in overall genome editing efficiency translate into better gene-targeting outcomes. Continued efforts to refine CRISPR/Cas-based systems, improve tissue-specific delivery, and optimize the design of repair templates are critical for maximizing HDR-mediated targeting efficiency. With time, these advancements are expected to expand the feasibility of HDR to more complex and larger sequence modifications. In parallel with these advances, alternative strategies that bypass the inherent limitations of HDR are emerging as powerful tools for achieving large sequence insertions.

TRANSPOSON-BASED DNA INSERTIONS

Transposable elements (TEs), or transposons, are natural DNA elements capable of moving throughout the genome (Mhiri et al., 2022). TEs are organized into two broad classes: type I and type II. Type I TEs move via an RNA intermediate in a 'copy-and-paste' manner, wherein the DNA element serves as a template for transcription. In contrast, type II TEs move via a DNA intermediate in a 'cut-and-paste' manner, in which the DNA element is excised from

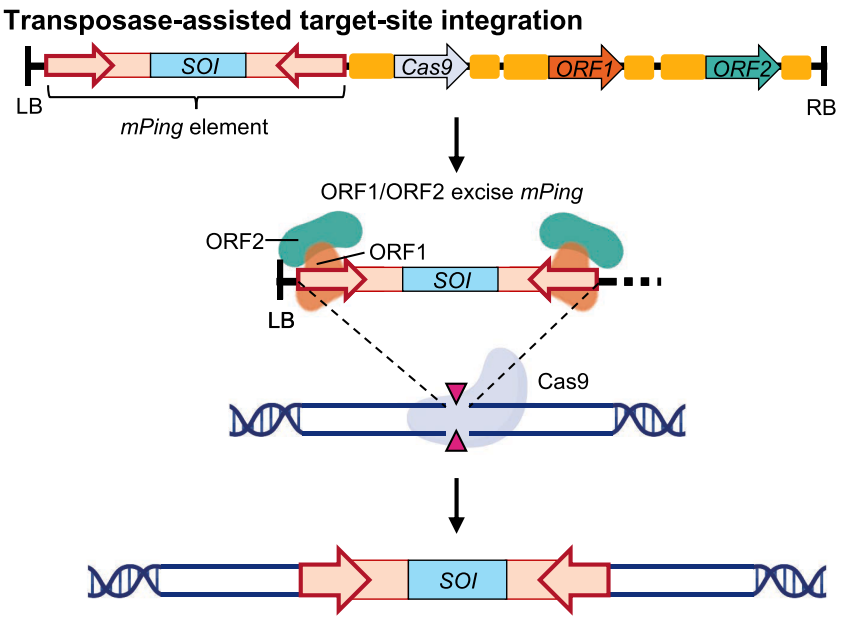

Figure 4. Transposon-based DNA insertions.
Transposase-assisted target-site integration (TATSI) uses the rice *mPing* transposable element in combination with Cas9 to mediate targeted transposon insertion. Cas9, gRNA, the *mPing* element, and transposase genes (*ORF1* and *ORF2*) are delivered on a T-DNA (promoters and terminators are represented by orange boxes). A sequence of interest (SOI) can be inserted within the *mPing* element for delivery into the host genome. ORF1 and ORF2 excise the mPing element from the T-DNA and direct insertion at the site of a double-stranded break, enabling Cas9-mediated delivery of a SOI.

one location in the genome and inserted into a new location (Mhiri et al., 2022). Historically, the 'cut-and-paste' nature of type II TEs has been widely used to move DNA elements somewhat randomly throughout the genome (Kirov, 2023). As most transposons exhibit a tendency for local transposition within the genome (Kunze & Weil, 2007; Moreno et al., 1992), some level of 'targeting' can be achieved by activating an inactive transposon near a desired target insertion site. By activating a previously inactive transposon, either by crossing in a transposase or transposase activation, one can introduce a transposon in proximity to a desired gene. This proximity effect has enabled the development of transposon insertional mutant libraries in various species, facilitating the identification and characterization of gene function (Brutnell, 2002).

A significant advancement in location-specific DNA integration has been the combination of Cas9-based technology with transposase/TE systems for transposase-assisted target-site integration (TATSI) in plants (Liu, Dong, et al., 2024). Liu, Dong, et al. (2024) designed a system that utilizes the rice *Ping/Pong* DNA transposon/transposase system combined with a programmable nuclease for targeted transposon insertion in Arabidopsis and soybean (Figure 4). These researchers tested various configurations of *SpCas9* fused to the *Pong* transposase genes (*ORF1* and *ORF2*), which are essential for *mPing* transposon excision and insertion. Fusing *Cas9* to the C-terminus of *ORF2* resulted in slightly reduced off-target integrations compared to all other fused and unfused configurations tested in this study (Liu, Dong, et al., 2024). Results revealed that 36% of T1 Arabidopsis plants contained the *mPing* TE insertion targeted upstream of the *ACTIN8* (*AtACT8*) gene, although deep sequencing still showed the occurrence of off-target integrations caused by free TE transposition. Liu et al. then tested the cargo capacity of the m*Ping* TE by engineering TE versions that contained additional DNA sequences for insertion. Insertions of up to 8.6 kb were obtained, although integration efficiency declined as cargo size increased. To demonstrate the commercial viability of the TATSI system, a *bar* (glyphosate resistance gene) transcriptional unit was inserted within the *mPing* TE and successfully targeted to a genome-safe-harbor site, *DD20*, in soybean (Liu, Dong, et al., 2024). Genome safe harbor sites are DNA sites that tolerate transgene insertion well due to the absence of coding genes, regulatory elements, miRNAs, lncRNAs, tRNAs, and centromeres (Sun et al., 2024). The TATSI system provides a relatively simple method to integrate custom DNA sequences in a location-specific manner with high efficiency. However, the TATSI system suffers from limitations, such as the introduction of over 400 nucleotides of transposon DNA along with the sequence of interest and a high rate of off-target integration. Future efforts aimed at reducing off-target effects and unintended transposon DNA integration are essential to ensure translation to commercial crop development (Liu, Dong, et al., 2024).

INTEGRASE-BASED DNA INSERTIONS

Bacterial enzymes known as site-specific recombinases (SSRs) are capable of DNA insertions, deletions, or inversions by recognizing and cleaving short, specific DNA sequences, following which SSRs exchange DNA fragments and rejoin the DNA strands (Smith, 2015). SSRs have been applied in plants for transgene marker removal, site-specific integration, and gene stacking (Ow, 2016). One challenge when applying SSRs in plants is that, unlike SSNs, SSRs are not programmable and only recognize long sequences (~20 nt) that must be pre-integrated into the genome as 'landing sites'. Only recently have researchers been able to precisely integrate recombinase landing sites at specific genomic loci using prime editing technology (Anzalone et al., 2022; Sun et al., 2024). Prime editing is a genome editing approach that uses a nickase Cas9 fused to a reverse transcriptase and a prime editing gRNA (pegRNA) to introduce small, targeted insertions, deletions, sequence replacements, and inversions (Anzalone et al., 2019; Molla et al., 2021; Zhao et al., 2025). Inspired by success in human cells (Anzalone et al., 2022; Yarnall et al., 2023), Sun et al. (2024) applied a tool coined as PrimeRoot for plant-optimized dual-enhanced prime editing combined with SSR technology to mediate large DNA insertions (up to 11.1 kb in rice) (Sun et al., 2024). The dual-enhanced prime editing system is used to introduce SSR sites in a specific DNA sequence. Two adjacent prime editing gRNAs (pegRNAs) are designed, where each is equipped with reverse transcription templates with the required sequence homology to introduce SSR target sites at a desired locus. A *SpCas9* nickase nicks the target site, and the nicked strand serves as a reverse transcription primer. Reverse transcriptase then introduces the desired changes using the pegRNA template. Since current prime editing tools are limited by size (an insertion of up to 300 bp has been achieved in plants) (Sun et al., 2024), instead, an SSR recombinase 'landing site' is integrated, facilitating larger targeted insertions at a given site. Following SSR landing site integration by prime editing, SSRs then excise the target sequence from the donor DNA and integrate it into the genomic recombinase sites, resulting in the precise incorporation of a desired sequence in the genome (Figure 5a). To test the capacity of this system, researchers tested DNA donors ranging from 1.4 to 11.1 kb in length and found only a slight decrease in efficiency with larger donor templates. To demonstrate some agriculturally relevant applications, Sun et al. (2024) introduced a gene cassette to confer rice blast resistance at a predicted genomic safe harbor site in the Kitaake rice cultivar, achieving 6.3% gene-targeting efficiency (Sun et al., 2024). PrimeRoot represents a novel method for generating targeted, large insertions that do not create DSBs. Some limitations include targeting restrictions imposed by dual-prime editing design specifications, recombination site 'scars,' and variable editing efficiency across sites (Sun et al., 2024).

Further optimization of prime editing and SSRs will likely continue to improve the efficiency and flexibility of this technique. One promising development is DNA-dependent DNA polymerase editing, an efficient technique to introduce modifications (>100 nt) implemented in mammalian cells (Figure 5b) (Liu, Panda, et al., 2024). DNA-dependent DNA polymerase editing uses an unfused nickase *SpCas9* to nick the target DNA strand. The nicked strand can then anneal with a synthetic linear DNA/RNA hybrid repair template through DNA base-pairing interactions. The repair template contains a 3′ RNA-based MS2 aptamer that is capable of binding a DNA polymerase-MCP fusion, bringing the high fidelity DNA polymerase (Phi29), repair template, and nicked target site in proximity to facilitate the desired sequence modification (Liu, Panda, et al., 2024). To demonstrate the application of DNA-dependent DNA polymerase editing, a single template was used to insert the 40 bp *loxP* site at the *ADENO-ASSOCIATED VIRUS INTEGRATION SITE1* with 35% efficiency. Importantly, DNA-dependent DNA polymerase is of higher fidelity and efficiency than the reverse transcriptase used in current prime editing approaches. Additionally, DNA polymerase editing templates can accommodate longer sequences than pegRNAs, which suffer from autoinhibitory intramolecular base pairing of the primer binding site with the pegRNA spacer (Liu, Panda, et al., 2024). This technique could be translated into plants as an alternative way to introduce small to medium modifications or SSR recognition sites using only a single template.

LOOKING FORWARD: DCAS9-SSAP, RECOMBINEERING, AND BRIDGERNAS

Plant biologists have long turned to innovations in other eukaryotic or prokaryotic systems to inspire tool development in plants. Three innovative approaches that will hopefully be translated into plants soon include a chimeric catalytically inactive Cas9 fusion with single-strand annealing protein (dCas9-SSAP), lambda red recombineering, and bridgeRNAs (Durrant et al., 2024; Hiraizumi et al., 2024; Thomason et al., 2023; Wang et al., 2022).

Single-strand annealing proteins (SSAPs) are bacteriophage-derived enzymes capable of recombination without induction of a DSB or nick. The association of a SSAP with dCas9 enabled programmable targeted DNA insertions without the creation of a DSB in mammalian cells (Wang et al., 2022). The RecT SSAP is recruited to dCas9 via a gRNA extension, i.e., an MS2 RNA aptamer that interacts with the MS2 coat protein (MCP) fused to RecT. The dCas9 is guided to and unwinds the target

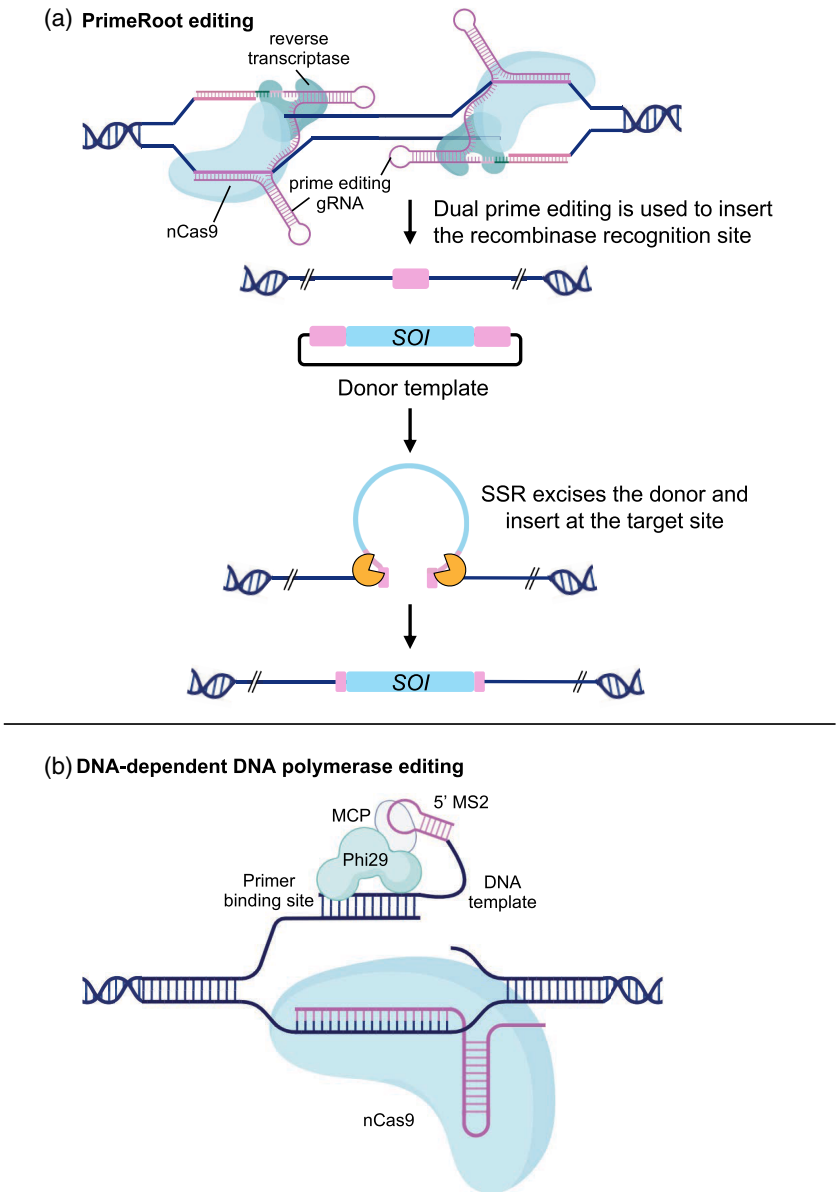

Figure 5. Integrase-based DNA insertions.
(a) PrimeRoot uses dual prime editing to introduce a recombinase recognition sequence at a desired sequence. The dual prime editing machinery consists of dual nickase Cas9 (nCas9), each equipped with prime editing gRNA and a fused reverse transcriptase. The nicked DNA strand acts as a reverse transcriptase primer, enabling the reverse transcriptase molecule to integrate the recombinase recognition sequence (part of the pegRNA) at the target locus. Then, the site-specific recombinase (SSR) excises a sequence of interest (SOI) from the donor template and directs recombination at the target recombinase integration site, resulting in a targeted insertion at a desired locus.
(b) DNA-dependent DNA polymerase editing is a technique that has been developed and employed in mammalian systems to introduce sequence insertions of up to 100 bp. A nickase Cas9 (nCas) creates a single-strand break at the target DNA site, generating a free DNA strand. A DNA–RNA hybrid template binds to this free stand via a DNA primer at the primer binding site. The DNA–RNA template is tethered to the MS2 coat protein (MCP) fusion with the DNA polymerase Phi29 through an RNA MS2 loop at the 5′ end of the DNA repair template. This interaction creates spatial proximity between the DNA polymerase, DNA template and nicked DNA target site, facilitating efficient and high-fidelity sequence insertion. Figure created in part using BioRender.

insertion site DNA, mediating R-loop formation, while the SSAP orchestrates homology-mediated integration of the donor DNA (Figure 6a). dCas9/SSAP successfully integrates large DNA fragments (>1 kb) with 5-fold higher editing efficiency than dCas9 alone, reaching up to 20% efficiency with no detected off-target insertions (Wang et al., 2022). The high precision, efficiency, and lack of off-target effects associated with dCas9/SSAP in mammalian cells are extremely encouraging, but it remains to be seen if SSAPs will function well in plant cells.

Recombineering is a homologous recombination-based gene engineering method implemented in bacteria. It relies on phage machinery expressed in an inducible manner in *E. coli* to edit DNA *in vivo* in the context of the bacterial genome or plasmid vector (Thomason et al., 2023). This technology can be employed to make targeted

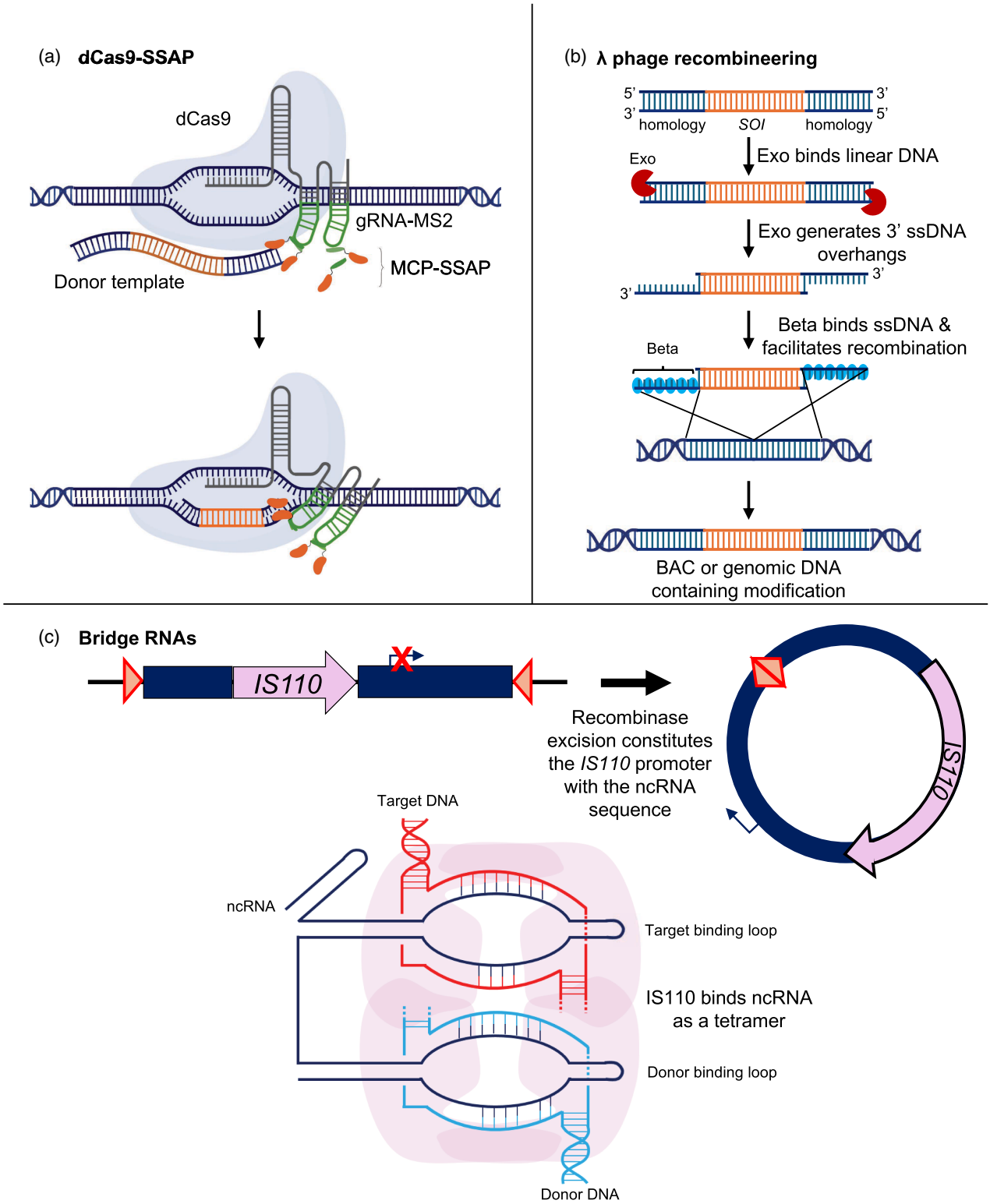

Figure 6. Emerging technologies.
(a) A nuclease-dead Cas9 (dCas9) fused with bacteriophage-derived single-strand annealing protein (SSAP) can mediate programmable DNA insertions in mammalian cells. A modified gRNA containing MS2 loops (light green) is capable of binding the SSAP-MS2 coat protein (MCP) fusion, bringing the SSAP in proximity to the desired editing site. dCas9 unwinds the target site DNA, making the region accessible for SSAP to orchestrate homology-mediated integration of a donor sequence (orange).
(b) λ phage recombineering makes use of three key proteins to mediate recombination in *E. coli*: Exo (red), Beta (cyan), and Gam (not pictured). A recombineering-competent strain of *E. coli* is transformed with a linear DNA fragment containing a sequence of interest (SOI) and short regions of homology with the target modification site. Exo binds linear DNA and generates 3′ ssDNA overhangs through 5′ to 3′ exonuclease activity. Beta proteins bind ssDNA, protecting it from degradation and facilitating recombination at the target locus.
(c) The *IS110* insertion sequence is a bacterial mobile genetic element that is capable of self-excision from its genomic context and assuming a dsDNA circular form (right). This circular form constitutes the IS110 promoter with a ncRNA sequence enabling its expression. This ncRNA (dark blue, below) serves as an RNA bridge: one of its structural loops binds both strands of the target DNA sequence (red) in the genome, and another binds both strands of the donor DNA (cyan), aligning the two DNA sites for recombination by the *IS110*-encoded recombinase (pink shape, bound as a tetramer to ncRNA). Figure created in part using BioRender.

deletions, insertions, inversions, and replacements of variable sizes, from single nucleotides to several kilobases. One such system, lambda red, relies on three lambda bacteriophage proteins, Exo, Bet, and Gam, expressed in *E. coli* in an inducible manner to catalyze the recombination between a user-provided linear DNA repair template (typically delivered into bacterial cells via electroporation) and the target locus (Figure 6b) (Murphy, 2016). Exo is an exonuclease that processes the dsDNA repair template to produce recombinogenic ssDNA ends, Bet is a SSAP that brings together and recombines the template with the target locus, and Gam keeps endogenous *E. coli* nucleases in check to prevent template degradation. As little as 40 bp of homology at the ends of the DNA repair template is sufficient for Bet to do its job. Recombineering has been successfully used to edit phage, bacterial, plant, and animal DNA in *E. coli* (Ayadi et al., 2012; Sarov et al., 2006; Zhang et al., 1998; Zhou et al., 2011), but to our knowledge, no studies have yet been published that succeeded in implementing recombineering directly in eukaryotes. To date, recombineering performed in *E. coli* large transformable BACs harboring Arabidopsis genomic DNA has been successful in making a variety of edits, from single-nucleotide replacements to multi-gene deletions to reporter integration (Bitrià n et al., 2011; Hu et al., 2019, 2022; Stepanova et al., 2008; Zhou et al., 2011). The *Agrobacterium*-mediated transformation of these constructs could successfully deliver and express in Arabidopsis the recombineered transgenes of up to 82 kb total cargo size (Stepanova et al., 2008; Zhou et al., 2011). What would be even more exciting, however, is to have the recombineering technology implemented *in planta* by integrating an inducible *Bet* recombinase and helper proteins directly into the plant genome. Although it is likely going to be challenging to get Bet to work effectively in plants due to the limited accessibility of chromatin-packaged genomic DNA, the results in mammalian cells using bacteriophage-derived SSAPs described above are encouraging.

BridgeRNAs are a programmable recombination tool developed in bacteria that will likely open a new frontier of genome editing across species. *IS110*-family insertion sequences are bacterial mobile genetic elements that, when excised from the genome, form a circular DNA structure capable of expressing a structural non-coding (nc) RNA and *IS110*-family recombinase (Durrant et al., 2024; Hiraizumi et al., 2024). This ncRNA binds to the IS110 recombinase and contains internal loops that base-pair with target and donor DNA, 'bridging' these DNA sites and enabling the recombinase-mediated seamless recombination (Figure 6c). Durrant et al. reprogrammed bridgeRNAs to direct recombinase activity in a plasmid-based recombination assay to seven different genomic targets with a 13.8–59.5% efficiency in *E. coli*. This demonstrates the technology's genome editing potential. Recombination was then demonstrated in the native *E. coli* genome, with 51.6% of insertions occurring at the target site and the remaining 48.4% of insertions integrating in other locations throughout the *E. coli* genome. Target specificity was further improved by increasing the number of base-pairing interactions between the *IS110*-associated ncRNA and the target loop upon the extension of the target sequence from 4 to 7 bp (Durrant et al., 2024). Given the novelty of this discovery, it remains to be seen if these recombinases can be applied in plants and other eukaryotes whose DNA is less accessible due to chromatin. However, given that other bacterial RNA-associated proteins work well in eukaryotic cells (Nekrasov et al., 2013), it is reasonable that after some optimization, these recombinases may facilitate programmable DNA insertion, excision, or inversion in plants.

REMAINING CHALLENGES

Despite the development of large DNA modification technologies, several challenges hinder the practical applications for model and crop plant species. Given that the current methods for gene insertion primarily rely on HDR-based techniques, they are often inefficient because they depend on the cell's natural DNA repair processes. This low efficiency makes it difficult to identify successful editing events, especially without the use of visual phenotypes or selectable markers. In fact, most of the technologies discussed herein that provided a proof-of-concept

demonstration of their editing capability leveraged phenotypic changes or included selectable markers in DNA cargos in gene targeting and insertion experiments. However, as gene targeting and sequence insertion technologies advance, one can imagine an expansion of DNA cargoes to include larger and more complex sequences, enabling precision gene insertion at safe harbor sites, multi-gene stacking, or even the introduction of entire biochemical pathways.

Besides limited editing efficiency, HDR-mediated gene insertions remain somewhat error-prone and can result in one-sided HDR events, instances of homologous recombination in which only one end of the repair template recombines at the target locus and the other end of the repair template is integrated at the target locus via NHEJ, resulting in indels (Puchta, 1998; Puchta & Fauser, 2014). Researchers have designed strategies to take advantage of this phenomenon in mammalian systems by using a single homologous stretch of donor DNA that is repaired using a combination of NHEJ and HDR repair (Suzuki et al., 2019). Perhaps further optimization of repair template design, delivery, and target site modification can increase the efficiency and fidelity of insertions and limit unintended partial HDR events.

One critical factor in further improving gene targeting is the form of the DNA repair template. Gene targeting approaches rely on diverse donor template formats, such as ssDNA, ribonucleoprotein (RNP) complexes, and dsODNs, many of which are incompatible with *Agrobacterium*-mediated transformation, necessitating alternative donor delivery methods such as biolistic delivery. In the future, the direct delivery of CRISPR-Cas RNP complexes and donor repair templates instead of DNA-based delivery systems can help mitigate concerns about random genomic integration of transgene constructs and off-target effects of Cas9 in gene targeting experiments, reducing the need for time-consuming backcrossing or resource-intensive screening following editing experiments (Metje-Sprink et al., 2019; Svitashev et al., 2016; Woo et al., 2015).

Furthermore, machine learning models may soon fill the gaps in gRNA and repair template design to increase the efficiency of HDR. Machine learning models are powerful tools for designing gRNAs (Chuai et al., 2018; Wang et al., 2020), predicting off-target effects (Listgarten et al., 2018) and forecasting editing outcomes (Chen et al., 2019; Shen et al., 2018). Aside from CRISPR/Cas9-triggered mutagenesis, O'Brien et al. developed a machine learning model to determine the optimal gene targeting strategy for HDR-mediated editing, enabling users to design HDR-mediated gene targets with 83% higher efficiency than traditionally designed targets (O'Brien et al., 2019). While most of these tools have been trained on experimental data from synthetic contexts or mammalian systems, findings from these tools can likely be translated to plant gene editing.

A recent shift in focus towards integrase-based DNA insertions promises a highly efficient method for sequence-specific integration (Sun et al., 2024). Current integrase-based methods in plants remain technically challenging for most research applications. Additionally, these methods require the integration of recombinase recognition sites at the target locus and result in remnant recombination site 'scars' following successful integration. The development of programmable recombinases, such as the IS110 family of recombinases, in plants may open up a new era of genome editing in which programmable, seamless integration of donor DNA at a given locus is achievable (Durrant et al., 2024; Hiraizumi et al., 2024).

Independent of genome editing tools, significant hurdles to technology transfer in agriculturally relevant crop species remain (Hua et al., 2019). Many agriculturally relevant crops are recalcitrant to traditional transformation methods or plant regeneration (Anjanappa & Gruissem, 2021). These limitations necessitate improvements in delivery methods to ensure that editing tools and donor DNA templates reach target cells efficiently.

Beyond delivery challenges, the identification of genome safe-harbor sites can enable gene stacking (Ceccon et al., 2020) and ensure stable transgene expression; however, there are few established methods to identify genome safe-harbor sites in crop species (Cantos et al., 2014; Sun et al., 2024). More robust computational pipelines and subsequent validation of genome safe-harbor sites are needed to support crop engineering efforts.

Finally, it is important to consider the regulatory landscape when designing a gene-targeting experiment. Small, 'cisgenic' modifications that consist of minor edits face much more lenient regulations in the United States (Ahmad et al., 2023). Transgenic crops, however, still face tight restrictions globally, often lengthening the time to market and increasing the cost of crop development (Ahmad et al., 2023). To mitigate some regulatory and environmental concerns associated with genetically modified crops, transgene containment technologies are of paramount importance (Stockdale & Millwood, 2023). Crops that contain larger sequence modifications are currently subject to the same regulations as traditional GM crops (Vora et al., 2023). Thus, prior to commercialization, a thorough assessment of transgene contamination risk is required. Several strategies have been developed to prevent outcrossing with wild relatives and native species, such as engineered male sterility, maternal inheritance, genome incompatibility, or delayed flowering (Stockdale & Millwood, 2023).

In summary, the tremendous progress made in genome engineering in the past few years has changed the

manner and scale of genome modifications that are possible. This review aimed to highlight a few of the key approaches to enhance the frequency of large modifications that have been tested in plants, as well as illuminate some of the groundbreaking strategies that have been developed in other organisms. Improved transformation and regeneration techniques, combined with the optimization of new and existing technologies, will continue to improve our ability to generate large, sequence-specific modifications in a range of plant species.

ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation: KV is a recipient of a Graduate Research Fellowship Program award 2023356574, and ANS and JMA are funded by research grants 1444561, 2327912, 1940829, and 1750006. We are grateful to Drs. Josh Strable and Marcela Rojas-Pierce for their critical feedback on this manuscript.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

REFERENCES

Abel, P.P., **Nelson, R.S.**, **De, B.**, **Hoffmann, N.**, **Rogers, S.G.**, **Fraley, R.T.** ***et al.*** (1986) Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene. *Science*, **232**, 738–743. Available from: https://doi.org/10.1126/science.3457472

Abel, S. & **Theologis, A.** (1994) Transient transformation of Arabidopsis leaf protoplasts: a versatile experimental system to study gene expression. *The Plant Journal*, **5**, 421–427. Available from: https://doi.org/10.1111/j.1365-313X.1994.00421.x

Ahmad, A., **Jamil, A.** & **Munawar, N.** (2023) GMOs or non-GMOs? The CRISPR conundrum. *Frontiers in Plant Science*, **14**, 1232938. Available from: https://doi.org/10.3389/fpls.2023.1232938

Ali, Z., **Shami, A.**, **Sedeek, K.**, **Kamel, R.**, **Alhabsi, A.**, **Tehseen, M.** ***et al.*** (2020) Fusion of the Cas9 endonuclease and the VirD2 relaxase facilitates homology-directed repair for precise genome engineering in rice. *Communications Biology*, **3**, 44. Available from: https://doi.org/10.1038/s42003-020-0768-9

Alonso, J.M. & **Stepanova, A.N.** (2014) Arabidopsis transformation with large bacterial artificial chromosomes. *Methods in Molecular Biology (Clifton, NJ)*, **1062**, 271–283. Available from: https://doi.org/10.1007/978-1-62703-580-4_15

Altpeter, F., **Baisakh, N.**, **Beachy, R.**, **Bock, R.**, **Capell, T.**, **Christou, P.** ***et al.*** (2005) Particle bombardment and the genetic enhancement of crops: myths and realities. *Molecular Breeding*, **15**, 305–327. Available from: https://doi.org/10.1007/s11032-004-8001-y

Anjanappa, R.B. & **Gruissem, W.** (2021) Current progress and challenges in crop genetic transformation. *Journal of Plant Physiology*, **261**, 153411. Available from: https://doi.org/10.1016/j.jplph.2021.153411

Anzalone, A.V., **Gao, X.D.**, **Podracky, C.J.**, **Nelson, A.T.**, **Koblan, L.W.**, **Raguram, A.** ***et al.*** (2022) Programmable deletion, replacement, integration and inversion of large DNA sequences with twin prime editing. *Nature Biotechnology*, **40**, 731–740. Available from: https://doi.org/10.1038/s41587-021-01133-w

Ayadi, A., **Birling, M.-C.**, **Bottomley, J.**, **Bussell, J.**, **Fuchs, H.**, **Fray, M.** ***et al.*** (2012) Mouse large-scale phenotyping initiatives: overview of the European mouse disease clinic (EUMODIC) and of the Wellcome Trust sanger institute mouse genetics project. *Mammalian Genome*, **23**, 600–610. Available from: https://doi.org/10.1007/s00335-012-9418-y

Back, G. & **Walther, D.** (2021) Identification of cis-regulatory motifs in first introns and the prediction of intron-mediated enhancement of gene expression in *Arabidopsis thaliana*. *BMC Genomics*, **22**, 390. Available from: https://doi.org/10.1186/s12864-021-07711-1

Baltes, N.J., **Gil-Humanes, J.**, **Cermak, T.**, **Atkins, P.A.** & **Voytas, D.F.** (2014) DNA replicons for plant genome engineering. *Plant Cell*, **26**, 151–163. Available from: https://doi.org/10.1105/tpc.113.119792

Barone, P., **Wu, E.**, **Lenderts, B.**, **Anand, A.**, **Gordon-Kamm, W.**, **Svitashev, S.** ***et al.*** (2020) Efficient gene targeting in maize using inducible CRISPR-Cas9 and marker-free donor template. *Molecular Plant*, **13**, 1219–1227. Available from: https://doi.org/10.1016/j.molp.2020.06.008

Bart, R., **Chern, M.**, **Park, C.-J.**, **Bartley, L.** & **Ronald, P.C.** (2006) A novel system for gene silencing using siRNAs in rice leaf and stem-derived protoplasts. *Plant Methods*, **2**, 13. Available from: https://doi.org/10.1186/1746-4811-2-13

Bhagtaney, L. & **Sundarrajan, P.** (2023) An overview of tools for genome editing: ZFNs, mega nucleases, and TALENs. In: Pandita, D. & Pandita, A. (Eds.) *CRISPR/Cas-mediated genome editing in plants*. Burlington: Apple Academic Press, pp. 37–64.

Bitrián, M., **Roodbarkelari, F.**, **Horváth, M.** & **Koncz, C.** (2011) BAC-recombineering for studying plant gene regulation: developmental control and cellular localization of SnRK1 kinase subunits. *The Plant Journal*, **65**, 829–842. Available from: https://doi.org/10.1111/j.1365-313X.2010.04462.x

Brutnell, T.P. (2002) Transposon tagging in maize. *Functional & Integrative Genomics*, **2**, 4–12. Available from: https://doi.org/10.1007/s10142-001-0044-0

Burris, K.P., **Dlugosz, E.M.**, **Collins, A.G.**, **Stewart, C.N.** & **Lenaghan, S.C.** (2016) Development of a rapid, low-cost protoplast transfection system for switchgrass (*Panicum virgatum* L.). *Plant Cell Reports*, **35**, 693–704. Available from: https://doi.org/10.1007/s00299-015-1913-7

Butler, N.M., **Baltes, N.J.**, **Voytas, D.F.** & **Douches, D.S.** (2016) Geminivirus-mediated genome editing in potato (*Solanum tuberosum* L.) using sequence-specific nucleases. *Frontiers in Plant Science*, **7**, 1045. Available from: https://doi.org/10.3389/fpls.2016.01045

Cantos, C., **Francisco, P.**, **Trijatmiko, K.R.**, **Slamet-Loedin, I.** & **Chadha-Mohanty, P.K.** (2014) Identification of "safe harbor" loci in indica rice genome by harnessing the property of zinc-finger nucleases to induce DNA damage and repair. *Frontiers in Plant Science*, **5**, 302. Available from: https://doi.org/10.3389/fpls.2014.00302

Ceccon, C.C., **Caverzan, A.**, **Margis, R.**, **Salvadori, J.R.** & **Grando, M.F.** (2020) Gene stacking as a strategy to confer characteristics of agronomic importance in plants by genetic engineering. *Ciência Rural*, **50**, e20190207. Available from: https://doi.org/10.1590/0103-8478cr20190207

Čermák, T., **Baltes, N.J.**, **Čegan, R.**, **Zhang, Y.** & **Voytas, D.F.** (2015) High-frequency, precise modification of the tomato genome. *Genome Biology*, **16**, 232. Available from: https://doi.org/10.1186/s13059-015-0796-9

Čermák, T., **Curtin, S.J.**, **Gil-Humanes, J.**, **Čegan, R.**, **Kono, T.J.Y.**, **Konečná, E.** ***et al.*** (2017) A multipurpose toolkit to enable advanced genome engineering in plants. *Plant Cell*, **29**, 1196–1217. Available from: https://doi.org/10.1105/tpc.16.00922

Chen, W., **McKenna, A.**, **Schreiber, J.**, **Haeussler, M.**, **Yin, Y.**, **Agarwal, V.** ***et al.*** (2019) Massively parallel profiling and predictive modeling of the outcomes of CRISPR/Cas9-mediated double-strand break repair. *Nucleic Acids Research*, **47**, 7989–8003. Available from: https://doi.org/10.1093/nar/gkz487

Chuai, G., **Ma, H.**, **Yan, J.**, **Chen, M.**, **Hong, N.**, **Xue, D.** ***et al.*** (2018) DeepCRISPR: optimized CRISPR guide RNA design by deep learning. *Genome Biology*, **19**, 80. Available from: https://doi.org/10.1186/s13059-018-1459-4

Dahan-Meir, T., **Filler-Hayut, S.**, **Melamed-Bessudo, C.**, **Bocobza, S.**, **Czosnek, H.**, **Aharoni, A.** ***et al.*** (2018) Efficient in planta gene targeting in tomato using geminiviral replicons and the CRISPR/Cas9 system. *The Plant Journal*, **95**, 5–16. Available from: https://doi.org/10.1111/tpj.13932

De Pater, S., **Klemann, B.J.P.M.** & **Hooykaas, P.J.J.** (2018) True gene-targeting events by CRISPR/Cas-induced DSB repair of the PPO locus with an ectopically integrated repair template. *Scientific Reports*, **8**, 3338. Available from: https://doi.org/10.1038/s41598-018-21697-z

Dickinson, L., **Yuan, W.**, **LeBlanc, C.**, **Thomson, G.**, **Wang, S.** & **Jacob, Y.** (2023) Regulation of gene editing using T-DNA concatenation. *Nature Plants*, **9**, 1398–1408. Available from: https://doi.org/10.1038/s41477-023-01495-w

Dixon, A.S., **Schwinn, M.K.**, **Hall, M.P.**, **Zimmerman, K.**, **Otto, P.**, **Lubben, T.H.** ***et al.*** (2016) NanoLuc complementation reporter optimized for accurate measurement of protein interactions in cells. *ACS Chemical Biology*, **11**, 400–408. Available from: https://doi.org/10.1021/acschembio.5b00753

Dumas, F., **Duckely, M.**, **Pelczar, P.**, **Van Gelder, P.** & **Hohn, B.** (2001) An *Agrobacterium* VirE2 channel for transferred-DNA transport into plant cells. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 485–490. Available from: https://doi.org/10.1073/pnas.98.2.485

Durrant, M.G., **Perry, N.T.**, **Pai, J.J.**, **Jangid, A.R.**, **Athukoralage, J.S.**, **Hiraizumi, M.** ***et al.*** (2024) Bridge RNAs direct programmable recombination of target and donor DNA. *Nature*, **630**, 984–993. Available from: https://doi.org/10.1038/s41586-024-07552-4

Eini, O., **Schumann, N.**, **Niessen, M.** & **Varrelmann, M.** (2022) Targeted mutagenesis in plants using beet curly top virus for efficient delivery of CRISPR/Cas12a components. *New Biotechnology*, **67**, 1–11. Available from: https://doi.org/10.1016/j.nbt.2021.12.002

Endo, M., **Mikami, M.** & **Toki, S.** (2016) Biallelic gene targeting in Rice. *Plant Physiology*, **170**, 667–677. Available from: https://doi.org/10.1104/pp.15.01663

Ercolano, M.R., **Ballvora, A.**, **Paal, J.**, **Steinbiss, H.-H.**, **Salamini, F.** & **Gebhardt, C.** (2004) Functional complementation analysis in potato via biolistic transformation with BAC large DNA fragments. *Molecular Breeding*, **13**, 15–22. Available from: https://doi.org/10.1023/B:MOLB.0000012326.01128.5d

Fauser, F., **Roth, N.**, **Pacher, M.**, **Ilg, G.**, **Sánchez-Fernández, R.**, **Biesgen, C.** ***et al.*** (2012) In planta gene targeting. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 7535–7540. Available from: https://doi.org/10.1073/pnas.1202191109

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

Filler-Hayut, S., **Kniazev, K.**, **Melamed-Bessudo, C.** & **Levy, A.A.** (2021) Targeted inter-homologs recombination in Arabidopsis euchromatin and heterochromatin. *International Journal of Molecular Sciences*, **22**, 12096. Available from: https://doi.org/10.3390/ijms222212096

Francis, K.E. & **Spiker, S.** (2005) Identification of *Arabidopsis thaliana* transformants without selection reveals a high occurrence of silenced T-DNA integrations. *The Plant Journal*, **41**, 464–477. Available from: https://doi.org/10.1111/j.1365-313X.2004.02312.x

Gang, H., **Liu, G.**, **Zhang, M.**, **Zhao, Y.**, **Jiang, J.** & **Chen, S.** (2019) Comprehensive characterization of T-DNA integration induced chromosomal rearrangement in a birch T-DNA mutant. *BMC Genomics*, **20**, 311. Available from: https://doi.org/10.1186/s12864-019-5636-y

Gardner, J.M. & **Jaspersen, S.L.** (2014) Manipulating the yeast genome: deletion, mutation, and tagging by PCR. In: Smith, J.S. & Burke, D.J. (Eds.) *Yeast genetics, methods in molecular biology*. New York, NY: Springer New York, pp. 45–78. Available from: https://doi.org/10.1007/978-1-4939-1363-3_5

Gasparis, S., **Kała, M.**, **Przyborowski, M.**, **Łyżnik, L.A.**, **Orczyk, W.** & **Nadolska-Orczyk, A.** (2018) A simple and efficient CRISPR/Cas9 platform for induction of single and multiple, heritable mutations in barley (*Hordeum vulgare* L.). *Plant Methods*, **14**, 111. Available from: https://doi.org/10.1186/s13007-018-0382-8

Gelvin, S.B. (2017) Integration of *Agrobacterium* T-DNA into the plant genome. *Annual Review of Genetics*, **51**, 195–217. Available from: https://doi.org/10.1146/annurev-genet-120215-035320

Gilbertson, R.L., **Sudarshana, M.**, **Jiang, H.**, **Rojas, M.R.** & **Lucas, W.J.** (2003) Limitations on Geminivirus genome size imposed by plasmodesmata and virus-encoded movement protein: insights into DNA trafficking. *Plant Cell*, **15**, 2578–2591. Available from: https://doi.org/10.1105/tpc.015057

Gil-Humanes, J., **Wang, Y.**, **Liang, Z.**, **Shan, Q.**, **Ozuna, C.V.**, **Sánchez-León, S.** ***et al.*** (2017) High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. *The Plant Journal*, **89**, 1251–1262. Available from: https://doi.org/10.1111/tpj.13446

Gong, W., **Zhou, Y.**, **Wang, R.**, **Wei, X.**, **Zhang, L.**, **Dai, Y.** ***et al.*** (2021) Analysis of T-DNA integration events in transgenic rice. *Journal of Plant Physiology*, **266**, 153527. Available from: https://doi.org/10.1016/j.jplph.2021.153527

Grützner, R., **Martin, P.**, **Horn, C.**, **Mortensen, S.**, **Cram, E.J.**, **Lee-Parsons, C.W.T.** ***et al.*** (2021) High-efficiency genome editing in plants mediated by a Cas9 gene containing multiple introns. *Plant Communications*, **2**, 100135. Available from: https://doi.org/10.1016/j.xplc.2020.100135

Hahn, F., **Eisenhut, M.**, **Mantegazza, O.** & **Weber, A.P.M.** (2018) Homology-directed repair of a defective glabrous gene in Arabidopsis with Cas9-based gene targeting. *Frontiers in Plant Science*, **9**, 424. Available from: https://doi.org/10.3389/fpls.2018.00424

Hamdan, M.F., **Mohd Noor, S.N.**, **Abd-Aziz, N.**, **Pua, T.-L.** & **Tan, B.C.** (2022) Green revolution to gene revolution: technological advances in agriculture to feed the world. *Plants*, **11**(10), 1297. Available from: https://doi.org/10.3390/plants11101297

Hamilton, C.M. (1997) A binary-BAC system for plant transformation with high-molecular-weight DNA. *Gene*, **200**, 107–116. Available from: https://doi.org/10.1016/S0378-1119(97)00388-0

Herrera-Estrella, L., **Depicker, A.**, **Van Montagu, M.** & **Schell, J.** (1983) Expression of chimaeric genes transferred into plant cells using a Ti-plasmid-derived vector. *Nature*, **303**, 209–213. Available from: https://doi.org/10.1038/303209a0

Hiraizumi, M., **Perry, N.T.**, **Durrant, M.G.**, **Soma, T.**, **Nagahata, N.**, **Okazaki, S.** ***et al.*** (2024) Structural mechanism of bridge RNA-guided recombination. *Nature*, **630**, 994–1002. Available from: https://doi.org/10.1038/s41586-024-07570-2

Hu, Z., **Ghosh, A.** & **Koncz, C.** (2022) Plant gene modification by BAC recombineering. In: Reisch, C.R. (Ed.) *Recombineering: methods and protocols*. New York, NY: Springer US, pp. 71–84. Available from: https://doi.org/10.1007/978-1-0716-2233-9_6

Hu, Z., **Ghosh, A.**, **Stolze, S.C.**, **Horváth, M.**, **Bai, B.**, **Schaefer, S.** ***et al.*** (2019) Gene modification by fast-track recombineering for cellular localization and isolation of components of plant protein complexes. *The Plant Journal*, **100**, 411–429. Available from: https://doi.org/10.1111/tpj.14450

Hua, K., **Zhang, J.**, **Botella, J.R.**, **Ma, C.**, **Kong, F.**, **Liu, B.** ***et al.*** (2019) Perspectives on the application of genome-editing Technologies in Crop Breeding. *Molecular Plant*, **12**, 1047–1059. Available from: https://doi.org/10.1016/j.molp.2019.06.009

Huang, T., **Armstrong, B.**, **Schindele, P.** & **Puchta, H.** (2021) Efficient gene targeting in *Nicotiana tabacum* using CRISPR/SaCas9 and temperature tolerant LbCas12a. *Plant Biotechnology Journal*, **19**(7), 1314–1324. Available from: https://doi.org/10.1111/pbi.13546

Hummel, A.W., **Chauhan, R.D.**, **Cermak, T.**, **Mutka, A.M.**, **Vijayaraghavan, A.**, **Boyher, A.** ***et al.*** (2018) Allele exchange at the EPSPS locus confers glyphosate tolerance in cassava. *Plant Biotechnology Journal*, **16**, 1275–1282. Available from: https://doi.org/10.1111/pbi.12868

Janssen, J.M., **Chen, X.**, **Liu, J.** & **Gonçalves, M.A.F.V.** (2019) The chromatin structure of CRISPR-Cas9 target DNA controls the balance between mutagenic and homology-directed gene-editing events. *Molecular Therapy--Nucleic Acids*, **16**, 141–154. Available from: https://doi.org/10.1016/j.omtn.2019.02.009

Jiang, W., **Zhou, H.**, **Bi, H.**, **Fromm, M.**, **Yang, B.** & **Weeks, D.P.** (2013) Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Research*, **41**, e188. Available from: https://doi.org/10.1093/nar/gkt780

Jogam, P., **Sandhya, D.**, **Alok, A.**, **Peddaboina, V.**, **Allini, V.R.** & **Zhang, B.** (2022) A review on CRISPR/Cas-based epigenetic regulation in plants. *International Journal of Biological Macromolecules*, **219**, 1261–1271. Available from: https://doi.org/10.1016/j.ijbiomac.2022.08.182

Jupe, F., **Rivkin, A.C.**, **Michael, T.P.**, **Zander, M.**, **Motley, S.T.**, **Sandoval, J.P.** ***et al.*** (2019) The complex architecture and epigenomic impact of plant T-DNA insertions. *PLoS Genetics*, **15**, e1007819. Available from: https://doi.org/10.1371/journal.pgen.1007819

Kim, J.H., **Yu, J.**, **Kim, H.K.**, **Kim, J.Y.**, **Kim, M.-S.**, **Cho, Y.-G.** ***et al.*** (2022) Genome editing of Golden SNP-carrying lycopene epsilon-cyclase (LcyE) gene using the CRSPR-Cas9/HDR and geminiviral replicon system in rice. *International Journal of Molecular Sciences*, **23**, 10383. Available from: https://doi.org/10.3390/ijms231810383

Kim, S.-I. & **Gelvin, S.B.** (2007) Genome-wide analysis of *Agrobacterium* T-DNA integration sites in the Arabidopsis genome generated under non-selective conditions. *The Plant Journal*, **51**, 779–791. Available from: https://doi.org/10.1111/j.1365-313X.2007.03183.x

Kirov, I. (2023) Toward transgene-free transposon-mediated biological mutagenesis for plant breeding. *International Journal of Molecular Sciences*, **24**, 17054. Available from: https://doi.org/10.3390/ijms242317054

Kiselev, K.V., Aleynova, O.A., Ogneva, Z.V., Suprun, A.R. & **Dubrovina, A.S.** (2021) 35S promoter-driven transgenes are variably expressed in different organs of *Arabidopsis thaliana* and in response to abiotic stress. *Molecular Biology Reports*, **48**, 2235–2241. Available from: https://doi.org/10.1007/s11033-021-06235-x

Kralemann, L.E.M., van Tol, N., Hooykaas, P.J.J. & **Tijsterman, M.** (2024) Molecular analysis of the role of polymerase theta in gene targeting in *Arabidopsis thaliana*. *The Plant Journal*, **118**, 255–262. Available from: https://doi.org/10.1111/tpj.16689

Kumar, J., Char, S.N., Weiss, T., Liu, H., Liu, B., Yang, B. ***et al.*** (2023) Efficient protein tagging and cis-regulatory element engineering via precise and directional oligonucleotide-based targeted insertion in plants. *Plant Cell*, **35**, 2722–2735. Available from: https://doi.org/10.1093/plcell/koad139

Kumar, S., Worden, A., Novak, S., Lee, R. & **Petolino, J.F.** (2016) A trait stacking system via intra-genomic homologous recombination. *Planta*, **244**, 1157–1166. Available from: https://doi.org/10.1007/s00425-016-2595-2

Kunze, R. & **Weil, C.F.** (2007) The hAT and CACTA superfamilies of plant transposons. In: Craig, N.L., Craigie, R., Gellert, M. & Lambowitz, A.M. (Eds.) *Mobile DNA II*. New York: Wiley, pp. 565–610. Available from: https://doi.org/10.1128/9781555817954.ch24

Kusano, H., Ohnuma, M., Mutsuro-Aoki, H., Asahi, T., Ichinosawa, D., Onodera, H. ***et al.*** (2018) Establishment of a modified CRISPR/Cas9 system with increased mutagenesis frequency using the translational enhancer dMac3 and multiple guide RNAs in potato. *Scientific Reports*, **8**, 13753. Available from: https://doi.org/10.1038/s41598-018-32049-2

Latham, J.R., Wilson, A.K. & **Steinbrecher, R.A.** (2006) The mutational consequences of plant transformation. *BioMed Research International*, **2006** (1), 025376. Available from: https://doi.org/10.1155/JBB/2006/25376

Lawrenson, T., Hinchliffe, A., Clarke, M., Morgan, Y. & **Harwood, W.** (2021) In-planta gene targeting in barley using Cas9 with and without geminiviral replicons. *Frontiers in Genome Editing*, **3**, 663380. Available from: https://doi.org/10.3389/fgeed.2021.663380

Li, B., Fu, C., Zhou, J., Hui, F., Wang, Q., Wang, F. ***et al.*** (2022) Highly efficient genome editing using Geminivirus-based CRISPR/Cas9 system in cotton plant. *Cells*, **11**, 2902. Available from: https://doi.org/10.3390/cells11182902

Li, G., Liu, D., Zhang, X., Quan, R., Zhong, C., Mo, J. ***et al.*** (2018) Suppressing Ku70/Ku80 expression elevates homology-directed repair efficiency in primary fibroblasts. *The International Journal of Biochemistry & Cell Biology*, **99**, 154–160. Available from: https://doi.org/10.1016/j.biocel.2018.04.011

Li, J., Meng, X., Zong, Y., Chen, K., Zhang, H., Liu, J. ***et al.*** (2016) Gene replacements and insertions in rice by intron targeting using CRISPR–Cas9. *Nature Plants*, **2**, 16139. Available from: https://doi.org/10.1038/nplants.2016.139

Li, S., Li, J., Zhang, J., Du, W., Fu, J., Sutar, S. ***et al.*** (2018) Synthesis-dependent repair of Cpf1-induced double strand DNA breaks enables targeted gene replacement in rice. *Journal of Experimental Botany*, **69**, 4715–4721. Available from: https://doi.org/10.1093/jxb/ery245

Li, S., Zhang, Y., Xia, L. & **Qi, Y.** (2020) CRISPR-Cas12a enables efficient biallelic gene targeting in rice. *Plant Biotechnology Journal*, **18**, 1351–1353. Available from: https://doi.org/10.1111/pbi.13295

Li, Z., Zhang, D., Xiong, X., Yan, B., Xie, W., Sheen, J. ***et al.*** (2017) A potent Cas9-derived gene activator for plant and mammalian cells. *Nature Plants*, **3**, 930–936. Available from: https://doi.org/10.1038/s41477-017-0046-0

Liang, Z., Chen, K., Li, T., Zhang, Y., Wang, Y., Zhao, Q. ***et al.*** (2017) Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. *Nature Communications*, **8**(1), 14261. Available from: https://doi.org/10.1038/ncomms14261

Listgarten, J., Weinstein, M., Kleinstiver, B.P., Sousa, A.A., Joung, J.K., Crawford, J. ***et al.*** (2018) Prediction of off-target activities for the end-to-end design of CRISPR guide RNAs. *Nature Biomedical Engineering*, **2**, 38–47. Available from: https://doi.org/10.1038/s41551-017-0178-6

Liu, B., Dong, X., Zheng, C., Keener, D., Chen, Z., Cheng, H. ***et al.*** (2024) Targeted genome editing with a DNA-dependent DNA polymerase and exogenous DNA-containing templates. *Nature Biotechnology*, **42**, 1039–1045. Available from: https://doi.org/10.1038/s41587-023-01947-w

Liu, J., Nannas, N.J., Fu, F., Shi, J., Aspinwall, B., Parrott, W.A. ***et al.*** (2019) Genome-scale sequence disruption following biolistic transformation in rice and maize. *Plant Cell*, **31**, 368–383. Available from: https://doi.org/10.1105/tpc.18.00613

Liu, P., Panda, K., Edwards, S.A., Swanson, R., Yi, H., Pandesha, P. ***et al.*** (2024) Transposase-assisted target-site integration for efficient plant genome engineering. *Nature*, **631**, 593–600. Available from: https://doi.org/10.1038/s41586-024-07613-8

Liu, W., Rudis, M.R., Cheplick, M.H., Millwood, R.J., Yang, J.-P., Ondzighi-Assoume, C.A. ***et al.*** (2020) Lipofection-mediated genome editing using DNA-free delivery of the Cas9/gRNA ribonucleoprotein into plant cells. *Plant Cell Reports*, **39**, 245–257. Available from: https://doi.org/10.1007/s00299-019-02488-w

Loyola-Vargas, V.M. & **Ochoa-Alejo, N.** (2018) An introduction to plant tissue culture: advances and perspectives. In: Loyola-Vargas, V.M. & Ochoa-Alejo, N. (Eds.) *Plant cell culture protocols*. New York, NY: Springer, pp. 3–13. Available from: https://doi.org/10.1007/978-1-4939-8594-4_1

Lu, Y., Tian, Y., Shen, R., Yao, Q., Wang, M., Chen, M. ***et al.*** (2020) Targeted, efficient sequence insertion and replacement in rice. *Nature Biotechnology*, **38**, 1402–1407. Available from: https://doi.org/10.1038/s41587-020-0581-5

Magembe, E.M., Li, H., Taheri, A., Zhou, S. & **Ghislain, M.** (2023) Identification of T-DNA structure and insertion site in transgenic crops using targeted capture sequencing. *Frontiers in Plant Science*, **14**, 665. Available from: https://doi.org/10.3389/fpls.2023.1156665

Malhi, G.S., Kaur, M. & **Kaushik, P.** (2021) Impact of climate change on agriculture and its mitigation strategies: a review. *Sustainability*, **13**(3), 1318. Available from: https://doi.org/10.3390/su13031318

Malzahn, A.A., Tang, X., Lee, K., Ren, Q., Sretenovic, S., Zhang, Y. ***et al.*** (2019) Application of CRISPR-Cas12a temperature sensitivity for improved genome editing in rice, maize, and Arabidopsis. *BMC Biology*, **17**(1), 9. Available from: https://doi.org/10.1186/s12915-019-0629-5

Mao, Y., Zhang, Z., Feng, Z., Wei, P., Zhang, H., Botella, J.R. ***et al.*** (2016) Development of germ-line-specific CRISPR-Cas9 systems to improve the production of heritable gene modifications in Arabidopsis. *Plant Biotechnology Journal*, **14**, 519–532. Available from: https://doi.org/10.1111/pbi.12468

Merker, L., Feller, L., Dorn, A. & **Puchta, H.** (2024) Deficiency of both classical and alternative end-joining pathways leads to a synergistic defect in double-strand break repair but not to an increase in homology-dependent gene targeting in Arabidopsis. *The Plant Journal*, **118**, 242–254. Available from: https://doi.org/10.1111/tpj.16604

Merker, L., Schindele, P. & **Puchta, H.** (2020) Using CRISPR/ttLbCas12a for in planta gene targeting in a. thaliana. *Current Protocols in Plant Biology*, **5**, e20117. Available from: https://doi.org/10.1002/cppb.20117

Metje-Sprink, J., Menz, J., Modrzejewski, D. & **Sprink, T.** (2019) DNA-free genome editing: past, present and future. *Frontiers in Plant Science*, **9**, 1957. Available from: https://doi.org/10.3389/fpls.2018.01957

Mhiri, C., Borges, F. & **Grandbastien, M.-A.** (2022) Specificities and dynamics of transposable elements in land plants. *Biology*, **11**, 488. Available from: https://doi.org/10.3390/biology11040488

Miki, D., Zhang, W., Zeng, W., Feng, Z. & **Zhu, J.-K.** (2018) CRISPR/Cas9-mediated gene targeting in Arabidopsis using sequential transformation. *Nature Communications*, **9**, 1967. Available from: https://doi.org/10.1038/s41467-018-04416-0

Molla, K.A., Sretenovic, S., Bansal, K.C. & **Qi, Y.** (2021) Precise plant genome editing using base editors and prime editors. *Nature Plants*, **7**, 1166–1187. Available from: https://doi.org/10.1038/s41477-021-00991-1

Molla, K.A. & **Yang, Y.** (2020) Predicting CRISPR/Cas9-induced mutations for precise genome editing. *Trends in Biotechnology*, **38**, 136–141. Available from: https://doi.org/10.1016/j.tibtech.2019.08.002

Moreno, M.A., Chen, J., Greenblatt, I. & **Dellaporta, S.L.** (1992) Reconstitutional mutagenesis of the maize P gene by short-range Ac transpositions. *Genetics*, **131**, 939–956. Available from: https://doi.org/10.1093/genetics/131.4.939

Movahedi, A., Wei, H., Zhou, X., Fountain, J.C., Chen, Z.-H., Mu, Z. ***et al.*** (2022) Precise exogenous insertion and sequence replacements in poplar by simultaneous HDR overexpression and NHEJ suppression using CRISPR-Cas9. *Horticulture Research*, **9**, uhac154. Available from: https://doi.org/10.1093/hr/uhac154

Murphy, K.C. (2016) λ Recombination and recombineering. *EcoSal Plus*, **7** (1), 3–8, 21–23. Available from: https://doi.org/10.1128/ecosalplus.ESP-0011–2015

Nagy, E.D., **Kuehn, R.**, **Wang, D.**, **Shrawat, A.**, **Duda, D.M.**, **Groat, J.R.** ***et al.*** (2022) Site-directed integration of exogenous DNA into the soybean genome by LbCas12a fused to a plant viral HUH endonuclease. *The Plant Journal*, **111**, 905–916. Available from: https://doi.org/10.1111/tpj.15849

Nekrasov, V., **Staskawicz, B.**, **Weigel, D.**, **Jones, J.D.G.** & **Kamoun, S.** (2013) Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. *Nature Biotechnology*, **31**, 691–693. Available from: https://doi.org/10.1038/nbt.2655

Nishizawa-Yokoi, A., **Nonaka, S.**, **Saika, H.**, **Kwon, Y.**, **Osakabe, K.** & **Toki, S.** (2012) Suppression of Ku70/80 or Lig4 leads to decreased stable transformation and enhanced homologous recombination in rice. *The New Phytologist*, **196**, 1048–1059. Available from: https://doi.org/10.1111/j.1469-8137.2012.04350.x

O'Brien, A.R., **Wilson, L.O.W.**, **Burgio, G.** & **Bauer, D.C.** (2019) Unlocking HDR-mediated nucleotide editing by identifying high-efficiency target sites using machine learning. *Scientific Reports*, **9**, 2788. Available from: https://doi.org/10.1038/s41598-019-39142-0

Olivares, F., **Loyola, R.**, **Olmedo, B.**, **Miccono, M.d.l.Á.**, **Aguirre, C.**, **Vergara, R.** ***et al.*** (2021) CRISPR/Cas9 targeted editing of genes associated with fungal susceptibility in *Vitis vinifera* L. cv. Thompson seedless using Geminivirus-derived replicons. *Frontiers in Plant Science*, **12**, 791030. Available from: https://doi.org/10.3389/fpls.2021.791030

O'Malley, R.C. & **Ecker, J.R.** (2010) Linking genotype to phenotype using the Arabidopsis unimutant collection. *Plant Journal: For Cell and Molecular Biology*, **61**, 928–940. Available from: https://doi.org/10.1111/j.1365-313X.2010.04119.x

Ow, D.W. (2016) The long road to recombinase-mediated plant transformation. *Plant Biotechnology Journal*, **14**, 441–447. Available from: https://doi.org/10.1111/pbi.12472

Partier, A., **Gay, G.**, **Tassy, C.**, **Beckert, M.**, **Feuillet, C.** & **Barret, P.** (2017) Molecular and FISH analyses of a 53-kbp intact DNA fragment inserted by biolistics in wheat (*Triticum aestivum* L.) genome. *Plant Cell Reports*, **36**, 1547–1559. Available from: https://doi.org/10.1007/s00299-017-2173-5

Peng, F., **Zhang, W.**, **Zeng, W.**, **Zhu, J.** & **Miki, D.** (2020) Gene targeting in Arabidopsis via an all-in-one strategy that uses a translational enhancer to aid Cas9 expression. *Plant Biotechnology Journal*, **18**, 892–894. Available from: https://doi.org/10.1111/pbi.13265

Phan, B.H., **Jin, W.**, **Topp, C.N.**, **Zhong, C.X.**, **Jiang, J.**, **Dawe, R.K.** ***et al.*** (2007) Transformation of rice with long DNA-segments consisting of random genomic DNA or centromere-specific DNA. *Transgenic Research*, **16**, 341–351. Available from: https://doi.org/10.1007/s11248-006-9041-3

Pingali, P.L. (2012) Green revolution: impacts, limits, and the path ahead. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 12302. Available from: https://doi.org/10.1073/pnas.0912953109

Puchta, H. (1998) Repair of genomic double-strand breaks in somatic plant cells by one-sided invasion of homologous sequences. *The Plant Journal*, **13**, 331–339. Available from: https://doi.org/10.1046/j.1365-313X.1998.00035.x

Puchta, H. (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *Journal of Experimental Botany*, **56**, 1–14. Available from: https://doi.org/10.1093/jxb/eri025

Puchta, H., **Dujon, B.** & **Hohn, B.** (1993) Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. *Nucleic Acids Research*, **21**, 5034–5040. Available from: https://doi.org/10.1093/nar/21.22.5034

Puchta, H. & **Fauser, F.** (2014) Synthetic nucleases for genome engineering in plants: prospects for a bright future. *Plant Journal: For Cell and Molecular Biology*, **78**, 727–741. Available from: https://doi.org/10.1111/tpj.12338

Pucker, B., **Kleinbölting, N.** & **Weisshaar, B.** (2021) Large scale genomic rearrangements in selected *Arabidopsis thaliana* T-DNA lines are caused by T-DNA insertion mutagenesis. *BMC Genomics*, **22**, 599. Available from: https://doi.org/10.1186/s12864-021-07877-8

Qi, Y., **Zhang, Y.**, **Zhang, F.**, **Baller, J.A.**, **Cleland, S.C.**, **Ryu, Y.** ***et al.*** (2013) Increasing frequencies of site-specific mutagenesis and gene targeting in Arabidopsis by manipulating DNA repair pathways. *Genome Research*, **23**, 547–554. Available from: https://doi.org/10.1101/gr.145557.112

Reed, K.M. & **Bargmann, B.O.R.** (2021) Protoplast regeneration and its use in new plant breeding technologies. *Frontiers in Genome Editing*, **3**, 734951. Available from: https://doi.org/10.3389/fgeed.2021.734951

Ren, R., **Gao, J.**, **Lu, C.**, **Wei, Y.**, **Jin, J.**, **Wong, S.-M.** ***et al.*** (2020) Highly efficient protoplast isolation and transient expression system for functional characterization of flowering related genes in cymbidium orchids. *International Journal of Molecular Sciences*, **21**, 2264. Available from: https://doi.org/10.3390/ijms21072264

Renaud, J.-B., **Boix, C.**, **Charpentier, M.**, **De Cian, A.**, **Cochennec, J.**, **Duvernois-Berthet, E.** ***et al.*** (2016) Improved genome editing efficiency and flexibility using modified oligonucleotides with TALEN and CRISPR-Cas9 nucleases. *Cell Reports*, **14**, 2263–2272. Available from: https://doi.org/10.1016/j.celrep.2016.02.018

Sanford, J.C. (1988) The biolistic process. *Trends in Biotechnology*, **6**, 299–302. Available from: https://doi.org/10.1016/0167-7799(88)90023-6

Sargent, R.G., **Brenneman, M.A.** & **Wilson, J.H.** (1997) Repair of site-specific double-strand breaks in a mammalian chromosome by homologous and illegitimate recombination. *Molecular and Cellular Biology*, **17**, 267–277. Available from: https://doi.org/10.1128/MCB.17.1.267

Sarov, M., **Schneider, S.**, **Pozniakovski, A.**, **Roguev, A.**, **Ernst, S.**, **Zhang, Y.** ***et al.*** (2006) A recombineering pipeline for functional genomics applied to *Caenorhabditis elegans*. *Nature Methods*, **3**, 839–844. Available from: https://doi.org/10.1038/nmeth933

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

Schindele, P., **Merker, L.**, **Schreiber, T.**, **Prange, A.**, **Tissier, A.** & **Puchta, H.** (2023) Enhancing gene editing and gene targeting efficiencies in *Arabidopsis thaliana* by using an intron-containing version of ttLbCas12a. *Plant Biotechnology Journal*, **21**, 457–459. Available from: https://doi.org/10.1111/pbi.13964

Schmidt, C., **Pacher, M.** & **Puchta, H.** (2019) Efficient induction of heritable inversions in plant genomes using the CRISPR/Cas system. *The Plant Journal*, **98**, 577–589. Available from: https://doi.org/10.1111/tpj.14322

Schreiber, T., **Prange, A.**, **Schäfer, P.**, **Iwen, T.**, **Grützner, R.**, **Marillonnet, S.** ***et al.*** (2024) Efficient scar-free knock-ins of several kilobases in plants by engineered CRISPR-Cas endonucleases. *Molecular Plant*, **17**, 824–837. Available from: https://doi.org/10.1016/j.molp.2024.03.013

Shaked, H., **Melamed-Bessudo, C.** & **Levy, A.A.** (2005) High-frequency gene targeting in Arabidopsis plants expressing the yeast RAD54 gene. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 12265–12269. Available from: https://doi.org/10.1073/pnas.0502601102

Shan, Q., **Baltes, N.J.**, **Atkins, P.**, **Kirkland, E.R.**, **Zhang, Y.**, **Baller, J.A.** ***et al.*** (2018) ZFN, TALEN and CRISPR-Cas9 mediated homology directed gene insertion in Arabidopsis: a disconnect between somatic and germinal cells. *Journal of Genetics and Genomics*, **45**, 681–684. Available from: https://doi.org/10.1016/j.jgg.2018.07.011

Sheen, J. (2001) Signal transduction in maize and Arabidopsis mesophyll protoplasts. *Plant Physiology*, **127**, 1466–1475. Available from: https://doi.org/10.1104/pp.010820

Shen, M.W., **Arbab, M.**, **Hsu, J.Y.**, **Worstell, D.**, **Culbertson, S.J.**, **Krabbe, O.** ***et al.*** (2018) Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature*, **563**, 646–651. Available from: https://doi.org/10.1038/s41586-018-0686-x

Shillito, R.D., **Saul, M.W.**, **Paszkowski, J.**, **Müller, M.** & **Potrykus, I.** (1985) High efficiency direct gene transfer to plants. *Bio/Technology*, **3**, 1099–1103. Available from: https://doi.org/10.1038/nbt1285-1099

Shilo, S., **Tripathi, P.**, **Melamed-Bessudo, C.**, **Tzfadia, O.**, **Muth, T.R.** & **Levy, A.A.** (2017) T-DNA-genome junctions form early after infection and are influenced by the chromatin state of the host genome. *PLoS Genetics*, **13**, e1006875. Available from: https://doi.org/10.1371/journal.pgen.1006875

Shou, J., **Li, J.**, **Liu, Y.** & **Wu, Q.** (2018) Precise and predictable CRISPR chromosomal rearrangements reveal principles of Cas9-mediated nucleotide insertion. *Molecular Cell*, **71**, 498–509. Available from: https://doi.org/10.1016/j.molcel.2018.06.021

Shukla, V.K., **Doyon, Y.**, **Miller, J.C.**, **DeKelver, R.C.**, **Moehle, E.A.**, **Worden, S.E.** ***et al.*** (2009) Precise genome modification in the crop species *Zea*

mays using zinc-finger nucleases. *Nature*, **459**(7245), 437–441. Available from: https://doi.org/10.1038/nature07992

Skarzyńska, A., **Pawełkowicz, M.** & **Pląder, W.** (2021) Influence of transgenesis on genome variability in cucumber lines with a thaumatin II gene. *Physiology and Molecular Biology of Plants*, **27**, 985–996. Available from: https://doi.org/10.1007/s12298-021-00990-8

Smith, M.C.M. (2015) Phage-encoded serine integrases and other large serine recombinases. In: Craig, N.L., Chandler, M., Gellert, M., Lambowitz, A., Rice, P.A. & Sandmeyer, S.B. (Eds.) *Mobile DNA III*. New York: John Wiley & Sons, Ltd, pp. 253–272. Available from: https://doi.org/10.1128/9781555819217.ch11

Squire, H.J., **Tomatz, S.**, **Voke, E.**, **González-Grandío, E.** & **Landry, M.** (2023) The emerging role of nanotechnology in plant genetic engineering. *Nature Reviews Bioengineering*, **1**, 314–328. Available from: https://doi.org/10.1038/s44222-023-00037-5

Stepanova, A.N., **Robertson-Hoyt, J.**, **Yun, J.**, **Benavente, L.M.**, **Xie, D.-Y.**, **Dolezal, K.** ***et al.*** (2008) TAA1-mediated auxin biosynthesis is essential for hormone crosstalk and plant development. *Cell*, **133**, 177–191. Available from: https://doi.org/10.1016/j.cell.2008.01.047

Stockdale, J.N. & **Millwood, R.J.** (2023) Transgene bioconfinement: don't flow there. *Plants*, **12**, 1099. Available from: https://doi.org/10.3390/plants12051099

Su, W., **Xu, M.**, **Radani, Y.** & **Yang, L.** (2023) Technological development and application of plant genetic transformation. *International Journal of Molecular Sciences*, **24**, 10646. Available from: https://doi.org/10.3390/ijms241310646

Sun, C., **Lei, Y.**, **Li, B.**, **Gao, Q.**, **Li, Y.**, **Cao, W.** ***et al.*** (2024) Precise integration of large DNA sequences in plant genomes using PrimeRoot editors. *Nature Biotechnology*, **42**, 316–327. Available from: https://doi.org/10.1038/s41587-023-01769-w

Sunilkumar, G., **Mohr, L.**, **Lopata-Finch, E.**, **Emani, C.** & **Rathore, K.S.** (2002) Developmental and tissue-specific expression of CaMV 35S promoter in cotton as revealed by GFP. *Plant Molecular Biology*, **50**, 463–479. Available from: https://doi.org/10.1023/A:1019832123444

Suzuki, K., **Yamamoto, M.**, **Hernandez-Benitez, R.**, **Li, Z.**, **Wei, C.**, **Soligalla, R.D.** ***et al.*** (2019) Precise in vivo genome editing via single homology arm donor mediated intron-targeting gene integration for genetic disease correction. *Cell Research*, **29**, 804–819. Available from: https://doi.org/10.1038/s41422-019-0213-0

Svitashev, S., **Schwartz, C.**, **Lenderts, B.**, **Young, J.K.** & **Mark Cigan, A.** (2016) Genome editing in maize directed by CRISPR–Cas9 ribonucleoprotein complexes. *Nature Communications*, **7**, 13274. Available from: https://doi.org/10.1038/ncomms13274

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

Tang, X., **Lowder, L.G.**, **Zhang, T.**, **Malzahn, A.A.**, **Zheng, X.**, **Voytas, D.F.** ***et al.*** (2017) A CRISPR–Cpf1 system for efficient genome editing and transcriptional repression in plants. *Nature Plants*, **3**, 17018. Available from: https://doi.org/10.1038/nplants.2017.18

Tang, Y., **Zhang, Z.**, **Yang, Z.** & **Wu, J.** (2023) CRISPR/Cas9 and *Agrobacterium tumefaciens* virulence proteins synergistically increase efficiency of precise genome editing via homology directed repair in plants. *Journal of Experimental Botany*, **74**, 3518–3530. Available from: https://doi.org/10.1093/jxb/erad096

Thomason, L.C., **Costantino, N.**, **Li, X.** & **Court, D.L.** (2023) Recombineering: genetic engineering in *Escherichia coli* using homologous recombination. *Current Protocols*, **3**, e656. Available from: https://doi.org/10.1002/cpz1.656

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

USDA, Economic Research Service (ERS). (2023) *More than half of harvested U.S. cropland uses seed varieties with at least one genetically modified trait*.

Vaeck, M., **Reynaerts, A.**, **Höfte, H.**, **Jansens, S.**, **De Beuckeleer, M.**, **Dean, C.** ***et al.*** (1987) Transgenic plants protected from insect attack. *Nature*, **328**, 33–37. Available from: https://doi.org/10.1038/328033a0

van Tol, N., **Schendel, R.**, **Bos, A.**, **van Kregten, M.**, **de Pater, S.**, **Hooykaas, P.J.J.** ***et al.*** (2022) Gene targeting in polymerase theta-deficient *Arabidopsis thaliana*. *The Plant Journal*, **109**, 112. Available from: https://doi.org/10.1111/tpj.15557

Vollen, K., **Zhao, C.**, **Alonso, J.M.** & **Stepanova, A.N.** (2024) Sourcing DNA parts for synthetic biology applications in plants. *Current Opinion in Biotechnology*, **87**, 103140. Available from: https://doi.org/10.1016/j.copbio.2024.103140

Vora, Z., **Pandya, J.**, **Sangh, C.** & **Vaikuntapu, P.R.** (2023) The evolving landscape of global regulations on genome-edited crops. *Journal of Plant Biochemistry and Biotechnology*, **32**, 831–845. Available from: https://doi.org/10.1007/s13562-023-00863-z

Vu, T.V., **Sivankalyani, V.**, **Kim, E.-J.**, **Doan, D.T.H.**, **Tran, M.T.**, **Kim, J.** ***et al.*** (2020) Highly efficient homology-directed repair using CRISPR/Cpf1-geminiviral replicon in tomato. *Plant Biotechnology Journal*, **18**, 2133–2143. Available from: https://doi.org/10.1111/pbi.13373

Wada, N., **Osakabe, K.** & **Osakabe, Y.** (2022) Expanding the plant genome editing toolbox with recently developed CRISPR–Cas systems. *Plant Physiology*, **188**, 1825–1837. Available from: https://doi.org/10.1093/plphys/kiac027

Wang, C., **Qu, Y.**, **Cheng, J.K.W.**, **Hughes, N.W.**, **Zhang, Q.**, **Wang, M.** ***et al.*** (2022) dCas9-based gene editing for cleavage-free genomic knock-in of long sequences. *Nature Cell Biology*, **24**, 268–278. Available from: https://doi.org/10.1038/s41556-021-00836-1

Wang, J., **Zhang, X.**, **Cheng, L.** & **Luo, Y.** (2020) An overview and metanalysis of machine and deep learning-based CRISPR gRNA design tools. *RNA Biology*, **17**, 13–22. Available from: https://doi.org/10.1080/15476286.2019.1669406

Wang, M., **Lu, Y.**, **Botella, J.R.**, **Mao, Y.**, **Hua, K.** & **Zhu, J.** (2017) Gene targeting by homology-directed repair in Rice using a Geminivirus-based CRISPR/Cas9 system. *Molecular Plant*, **10**, 1007–1010. Available from: https://doi.org/10.1016/j.molp.2017.03.002

Wang, Y., **Cheng, X.**, **Shan, Q.**, **Zhang, Y.**, **Liu, J.**, **Gao, C.** ***et al.*** (2014) Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nature Biotechnology*, **32**, 947–951. Available from: https://doi.org/10.1038/nbt.2969

Wang, Z.-P., **Xing, H.-L.**, **Dong, L.**, **Zhang, H.-Y.**, **Han, C.-Y.**, **Wang, X.-C.** ***et al.*** (2015) Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. *Genome Biology*, **16**, 144. Available from: https://doi.org/10.1186/s13059-015-0715-0

Weising, K., **Bohn, H.** & **Kahl, G.** (1990) Chromatin structure of transferred genes in transgenic plants. *Developmental Genetics*, **11**, 233–247. Available from: https://doi.org/10.1002/dvg.1020110309

Wolter, F., **Klemm, J.** & **Puchta, H.** (2018) Efficient in planta gene targeting in Arabidopsis using egg cell-specific expression of the Cas9 nuclease of *Staphylococcus aureus*. *The Plant Journal*, **94**, 735–746. Available from: https://doi.org/10.1111/tpj.13893

Wolter, F. & **Puchta, H.** (2019) In planta gene targeting can be enhanced by the use of CRISPR/Cas12a. *The Plant Journal*, **100**, 1083–1094. Available from: https://doi.org/10.1111/tpj.14488

Woo, J.W., **Kim, J.**, **Kwon, S.I.**, **Corvalán, C.**, **Cho, S.W.**, **Kim, H.** ***et al.*** (2015) DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nature Biotechnology*, **33**, 1162–1164. Available from: https://doi.org/10.1038/nbt.3389

Wright, D.A., **Townsend, J.A.**, **Winfrey, R.J., Jr.**, **Irwin, P.A.**, **Rajagopal, J.**, **Lonosky, P.M.** ***et al.*** (2005) High-frequency homologous recombination in plants mediated by zinc-finger nucleases. *The Plant Journal*, **44**, 693–705. Available from: https://doi.org/10.1111/j.1365-313X.2005.02551.x

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

Yan, L., **Wei, S.**, **Wu, Y.**, **Hu, R.**, **Li, H.**, **Yang, W.** ***et al.*** (2015) High-efficiency genome editing in Arabidopsis using YAO promoter-driven CRISPR/Cas9 system. *Molecular Plant*, **8**, 1820–1823. Available from: https://doi.org/10.1016/j.molp.2015.10.004

Yarnall, M.T.N., **Ioannidi, E.I.**, **Schmitt-Ulms, C.**, **Krajeski, R.N.**, **Lim, J.**, **Villiger, L.** ***et al.*** (2023) Drag-and-drop genome insertion of large sequences without double-strand DNA cleavage using CRISPR-directed integrases.

Nature Biotechnology, **41**, 500–512. Available from: https://doi.org/10.1038/s41587-022-01527-4

Zhang, F., **Maeder, M.L.**, **Unger-Wallace, E.**, **Hoshaw, J.P.**, **Reyon, D.**, **Christian, M.** ***et al.*** (2010) High frequency targeted mutagenesis in *Arabidopsis thaliana* using zinc finger nucleases. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 12028–12033. Available from: https://doi.org/10.1073/pnas.0914991107

Zhang, K., **Liu, S.**, **Fu, Y.**, **Wang, Z.**, **Yang, X.**, **Li, W.** ***et al.*** (2023) Establishment of an efficient cotton root protoplast isolation protocol suitable for single-cell RNA sequencing and transient gene expression analysis. *Plant Methods*, **19**(1), 5. Available from: https://doi.org/10.1186/s13007-023-00983-6

Zhang, X., **Henriques, R.**, **Lin, S.-S.**, **Niu, Q.-W.** & **Chua, N.-H.** (2006) *Agrobacterium*-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nature Protocols*, **1**, 641–646. Available from: https://doi.org/10.1038/nprot.2006.97

Zhang, Y., **Buchholz, F.**, **Muyrers, J.P.P.** & **Stewart, A.F.** (1998) A new logic for DNA engineering using recombination in *Escherichia coli*. *Nature Genetics*, **20**, 123–128. Available from: https://doi.org/10.1038/2417

Zhang, Y., **Malzahn, A.A.**, **Sretenovic, S.** & **Qi, Y.** (2019) The emerging and uncultivated potential of CRISPR technology in plant science. *Nature Plants*, **5**, 778–794. Available from: https://doi.org/10.1038/s41477-019-0461-5

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

Zhao, Y., **Huang, Z.**, **Zhou, X.**, **Teng, W.**, **Liu, Z.**, **Wang, W.** ***et al.*** (2025) Precise deletion, replacement and inversion of large DNA fragments in plants using dual prime editing. *Nature Plants*, **11**, 191–205. Available from: https://doi.org/10.1038/s41477-024-01898-3

Zhao, Z., **Shang, P.**, **Sage, F.** & **Geijsen, N.** (2022) Ligation-assisted homologous recombination enables precise genome editing by deploying both MMEJ and HDR. *Nucleic Acids Research*, **50**, e62. Available from: https://doi.org/10.1093/nar/gkac118

Zhou, R., **Benavente, L.M.**, **Stepanova, A.N.** & **Alonso, J.M.** (2011) A recombineering-based gene tagging system for Arabidopsis. *The Plant Journal*, **66**, 712–723. Available from: https://doi.org/10.1111/j.1365-313X.2011.04524.x

Zhu, H., **Li, C.** & **Gao, C.** (2020) Applications of CRISPR–Cas in agriculture and plant biotechnology. *Nature Reviews. Molecular Cell Biology*, **21**, 661–677. Available from: https://doi.org/10.1038/s41580-020-00288-9

Anzalone, A.V., **Randolph, P.B.**, **Davis, J.R.**, **Sousa, A.A.**, **Koblan, L.W.**, **Levy, J.M.** (2019) Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature,* **576**, 149–157. https://doi.org/10.1038/s41586-019-1711-4