




Review

# Plant DNA Repair and *Agrobacterium* T–DNA Integration

Stanton B. Gelvin 

Department of Biological Sciences, Purdue University, West Lafayette, IN 47907-1392, USA; gelvin@purdue.edu

**Abstract:** *Agrobacterium* species transfer DNA (T–DNA) to plant cells where it may integrate into plant chromosomes. The process of integration is thought to involve invasion and ligation of T–DNA, or its copying, into nicks or breaks in the host genome. Integrated T–DNA often contains, at its junctions with plant DNA, deletions of T–DNA or plant DNA, filler DNA, and/or microhomology between T–DNA and plant DNA pre-integration sites. T–DNA integration is also often associated with major plant genome rearrangements, including inversions and translocations. These characteristics are similar to those often found after repair of DNA breaks, and thus DNA repair mechanisms have frequently been invoked to explain the mechanism of T–DNA integration. However, the involvement of specific plant DNA repair proteins and *Agrobacterium* proteins in integration remains controversial, with numerous contradictory results reported in the literature. In this review I discuss this literature and comment on many of these studies. I conclude that either multiple known DNA repair pathways can be used for integration, or that some yet unknown pathway must exist to facilitate T–DNA integration into the plant genome.

**Keywords:** *Agrobacterium*; chromatin; DNA polymerase  $\theta$ ; DNA repair; genome rearrangements; microhomology–mediated end–joining (MMEJ); non-homologous end-joining (NHEJ); T–DNA borders; T–DNA integration



**Citation:** Gelvin, S.B. Plant DNA Repair and *Agrobacterium* T–DNA Integration. *Int. J. Mol. Sci.* **2021**, *22*, 8458. <https://doi.org/10.3390/ijms22168458>

Academic Editor: Jean Molinier

Received: 15 July 2021

Accepted: 3 August 2021

Published: 6 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The process of *Agrobacterium* T-DNA integration into the genomes of infected plants has important implications both for understanding plant DNA break repair processes and for the use of *Agrobacterium* as a tool for manipulating the plant genome. Although scientists have studied T-DNA integration for decades, we still have a very incomplete picture of how integration occurs, and numerous contradictory conclusions abound in the literature. In this short article, I summarize and analyze some of these conclusions, and relate what we know about plant DNA repair processes to possible mechanisms of T-DNA integration. For past discussions regarding the mechanism of T-DNA integration, I refer the reader to [1–10].

## 2. Are *Agrobacterium* Proteins Involved in T-DNA Integration into the Plant Genome?

The pathway of T-DNA transfer from *Agrobacterium* to plant cells, and its ultimate integration into the plant genome, starts with nicking the T-DNA region of a Ti (tumor inducing) or Ri (rhizogenic) plasmid by the T-DNA border-specific endonuclease VirD2 [11–15]. Nicking occurs between nucleotides 3 and 4 of the 25 bp border sequences that flank the T-DNA region [16,17]. During T-DNA border nicking, VirD2 covalently links to the 5' end of T-DNA, resulting in a single-strand form of T-DNA, the T-strand, that on its 3' end contains nucleotides 4–25 of the left border (LB), and on its 5' end contains nucleotides 1–3 of the right border (RB) [18–22]. VirD2 subsequently leads the T-strand through a dedicated type IV protein secretion system (T4SS) and into the plant cell [23,24]. Within the plant, T-strands may suffer deletions at the 3' and/or 5' ends before or during integration. Deletions are especially prevalent, and generally more extensive, at the 3' end than at the 5' end, which is protected by its linkage to VirD2 protein (e.g., [21]).

Does VirD2 directly participate in the T-DNA integration process? VirD2 can both cleave and re-ligate (i.e., reverse the reaction) T-DNA border sequences *in vitro* [25]. However, VirD2 does not harbor an activity that can ligate VirD2/T-strands to generalized target sequences; this could only be done by a ligase activity found in plant extracts [26].

VirD2 contains a highly conserved region (amino acids DGRGG) near the C-terminus, termed the  $\omega$  domain [27]. Substitution of DDGR (the first D is not part of  $\omega$ ) by four serine residues resulted in a mutant VirD2 protein that conferred somewhat lower transient transformation activity (~20–30% of wild-type levels) upon its *Agrobacterium* host compared with wild-type VirD2. However, stable transformation was reduced by >95% [27–29]. Although stable transformation of plants using this VirD2  $\omega$  mutant was decreased, the precision of the integration of sequences near the RB was similar to that observed when using wild-type VirD2 [29]. Mutation of other sequences in VirD2 protein, however, could alter the precision of T-DNA integrated near the RB [30]. Taken together, the results of these two studies suggest that VirD2 may be involved in T-DNA integration.

Although the VirD2  $\omega$  mutant *Agrobacterium* strain identified by Shurvinton et al. [27] showed moderately lower transient transformation activity but greatly reduced stable transformation, two other  $\omega$  domain mutants (a precise deletion of the DGRGG  $\omega$  amino acids, or their replacement with five glycine residues) resulted in both greatly decreased transient and stable transformation frequency [31]. The different relative transformation activities conferred by the various VirD2  $\omega$  mutants may result from altered protein structure conferred by the serine residue substitutions.

### 3. Where in the Plant Genome Does T-DNA Integrate?

Early studies indicated that T-DNA integration is random at the chromosome level [32–34]. Generation of numerous *Arabidopsis* and rice T-DNA insertion libraries, each consisting of tens of thousands of individually tagged plant genomes, allowed the first large scale probing of T-DNA insertion locations at the DNA sequence level. Results of these studies initially indicated that T-DNA preferentially integrated into transcriptionally active genes, promoter regions, or sequences of high A+T content [35–40]. These studies, however, all suffered from the problem of selection bias; the individual transgenic plants each harboring a different T-DNA integration event had been selected for antibiotic/herbicide resistance. If T-DNA had integrated into a transcriptionally inert region of the plant genome, the selection marker gene would not have been expressed and the resulting transgenic plant would have been lost. Indeed, fewer T-DNA insertions into heterochromatic regions of DNA, centromeres, telomeres, and rRNA genes were recovered relative to the proportion of the genomes represented by these sequences, resulting in the appearance of T-DNA integration into only transcriptionally active regions of the genome [37].

In contrast to the studies cited above, two groups examined T-DNA integration sites in the *Arabidopsis* genome in which cells were not selected for expression of any transgene, including those for antibiotic/herbicide resistance [41,42]. These experiments indicated that T-DNA did not preferentially integrate into any particular sequence context or region of gene expression. For example, approximately 10% of the *Arabidopsis* genome is composed of highly repeated DNA sequences, and ~10% of the insertions occurred in these sequences. T-DNA insertions into rDNA, centromeres, and telomeres also approximated their relative proportion of the genome. T-DNA pre-integration sites were average in their extent of transcription and methylation, as compared with the entire genome [41]. Thus, T-DNA integration did not preferentially occur into any particular chromosome sequence or feature. These results were reproduced using a high-throughput sequencing strategy to identify, without selection, T-DNA integrated into the *Arabidopsis* genome within six hours after infection [42]. This group also failed to identify preferential T-DNA integration into particular sequences, and the extent of DNA methylation of pre-integration sites was not biased. They did identify a slight local A+T motif enrichment at the pre-integration site, and microhomology was often observed between the T-DNA border sequences and the pre-integration site. Microhomology between T-DNA border regions and pre-integration

sites is a common feature of T-DNA integration (e.g., [43–45]). The one distinct feature of pre-integration sites was a high nucleosome occupancy and a high level of histone H2K27 trimethylation. However, data for these last two items were based on database entries and not on direct analysis of chromatin from the tissues the authors used in their studies.

Taken together, these two studies do not point to any distinctive plant DNA or chromatin features as T-DNA target integration sites. However, chromatin conformation may influence T-DNA integration. The *rat5 Arabidopsis* mutant is highly susceptible to transient transformation but highly resistant to stable transformation by *Agrobacterium* [46]. This mutant contains a T-DNA insertion in the 3' untranslated region of the histone H2A-1 gene *HTA1* [47]. Overexpression of *HTA1* in otherwise wild-type *Arabidopsis* plants increased stable transformation, as it also did in rice [48]. The expression of *HTA1* correlates with *Arabidopsis* cell and tissue susceptibility to *Agrobacterium*-mediated transformation [49], and all tested histone H2A proteins were functionally redundant with respect to increasing transformation when expressed in *Arabidopsis* [50]. In addition to histone H2A, overexpression of histone H4 and one histone H3 protein (HTR11) also conferred hyper-susceptibility to *Agrobacterium*-mediated transformation. However, overexpression of other histone H3 proteins and all tested histone H2B genes had no effect on transformation [51]. It is not clear whether manipulation of histone levels in plant cells directly alters the extent of T-DNA integration, or whether these altered histone levels influence plant gene expression, thereby resulting in T-DNA integration [52]. Nevertheless, the direct interaction of VirD2 with histones in yeast suggests that histones may help direct VirD2/T-strand complexes to the host genome [53].

Do *Agrobacterium* or host proteins guide T-strands to host chromatin prior to integration? VirE2 protein, a virulence effector protein secreted by *Agrobacterium* into plant cells, is a single-strand DNA-binding protein that has been proposed to interact with VirD2/T-strands in the plant cell, forming a “T-complex” [54]. (It should be noted that such complexes have never been identified in *Agrobacterium*-infected plants). VirE2 can interact with the plant bZIP transcription factor VIP1 [55]. VIP1 can interact with histones and nucleosomes, suggesting that VIP1 may be a molecular link between T-complexes entering the nucleus and T-DNA integration sites in the plant chromosomes [56–58]. However, two recent publications demonstrated that neither VIP1 nor its orthologs are required for *Agrobacterium*-mediated transformation, throwing into doubt a role for VIP1 in T-DNA integration [59,60].

In addition to the involvement of histones in *Agrobacterium*-mediated transformation and, perhaps, T-DNA integration, histone-associated and -modifying proteins also affect transformation. Crane and Gelvin [61] tested RNAi lines individually directed against 109 *Arabidopsis* chromatin-related genes for transient and stable transformation. Silencing of 24 of these genes decreased transformation. In particular, silencing of *SGA1* (encoding a histone H3 chaperone) and *HDT1* and *HDT2* (encoding histone deacetylases) greatly decreased both stable transformation and T-DNA integration. Silencing of genes involved in chromatin remodeling, DNA methylation, histone acetylation, and nucleosome assembly also had an effect on stable transformation, although this may result from secondary effects these genes have on the expression of other genes involved in stable transformation. Deletion of genes encoding components of yeast histone acetyltransferase complexes increased yeast transformation, whereas deletion of genes encoding proteins of histone deacetylase complexes decreased yeast transformation [62]. For some of these mutants, integration of T-DNA into the yeast genome was disrupted.

Taken together, these results indicate that proteins associated with chromatin structure and modification are important for T-DNA integration into plant or yeast genomes, and that in some instances, the effect on integration may be direct rather than indirectly influencing the expression of other genes important for T-DNA integration.

As described above, the position of T-DNA integration, and the chromatin structure of that region, may influence the expression of T-DNA-encoded transgenes. As a practical consideration, this variability in transgene expression (the so-called “position effect”) will

influence studies on gene and promoter function. Scientists therefore need to examine a large number of independent transgenic events to draw conclusions about, e.g., relative promoter strengths.

#### 4. Give Me a Break (?)

Integration of naked DNA can be increased by low-dose X-irradiation of transfected plant protoplasts, suggesting that generation of DNA breaks enhances foreign DNA integration [63]. Such integration events frequently result in deletions and rearrangements at the integration site. Furthermore, filler DNA from either within the introduced DNA or from elsewhere in the plant genome often appears at plasmid–plasmid junction sites, as does the presence of microhomology between molecules at the junction site [64]. Such genomic rearrangements are reminiscent of those caused by radiation-induced damage in *Arabidopsis* [65], suggesting that naked DNA integration into the plant genome occurs at DNA double–strand break sites. Similar rearrangements are frequently detected at junctions between integrated T-DNA and plant DNA, and between T-DNA borders of T-circles (see below). Thus, integration of various forms of exogenous DNA likely occurs by similar mechanisms, regardless of whether the introduced DNA is in a “naked” double-strand form or in the form of single-strand T-complexes after transfer from *Agrobacterium*. DNA break repair models have been invoked to understand how each of these forms of DNA integrates into the plant genome [66].

DNA repair and recombination processes require breaks in the phosphodiester DNA backbone, either in the form of single-strand nicks or double-strand breaks. Numerous models for T-DNA integration have been proposed; all these models incorporate, as part of their mechanism, nicks or breaks in the host target DNA at the site of T-DNA integration (see, e.g., [1,3–8,10]).

T-DNA preferentially integrates into double-strand DNA breaks [67]. This observation was followed by two other reports also showing preferential T-DNA integration into double-strand break sites [68,69]. In each of these studies, a rare cutting meganuclease (either I-*SceI* or I-*CeuI*) was used to cut tobacco DNA during transformation. T-DNA was preferentially “trapped” in these cut sites at frequencies up to several percent of the examined integration events. More recently, scientists used CRISPR technology to generate double-strand breaks in DNA, either to generate site-directed mutations or to attempt homology-dependent repair using recombination with correction templates. In several instances, T-DNA was trapped at these break sites following Cas nuclease cutting (e.g., [70,71]). It is thus clear that double-strand DNA breaks can act as a “T-DNA magnet”. However, does *Agrobacterium* take advantage of naturally occurring host DNA breaks (or nicks), or can *Agrobacterium* infection perhaps induce host DNA disruptions?

That *Agrobacterium* can incite DNA breaks would not be unusual, because inoculation by other plant pathogens (bacteria, oomycetes, and fungi) can cause double-strand DNA breaks in host plant genomes [72]. DNA disruptions occur in *Arabidopsis* cells near the site of *Agrobacterium* infection, as detected by COMET assays. However, because alkaline pH conditions were used in this study, it is not clear whether these disruptions resulted from single-strand nicks or double-strand breaks in the plant DNA [73]. Recent results indicate that *Arabidopsis* cells, exposed to *Agrobacterium* but not stably transformed, contain a higher number of in/dels than would be expected from the natural frequency of such mutations [74]. These results suggest that incubation of cells with *Agrobacterium* is inherently mutagenic, causing double-strand DNA breaks that are mis-repaired.

There are many hints in the literature that *Agrobacterium* infection can cause mutations independent of T-DNA integration; these mutations may result from induced double-strand DNA breaks that are subsequently mis-repaired. They may also be generated by “abortive integration” of T-DNA, followed by mis-repair of the abortive integration site. For example, *N. plumbaginifolia* plants, containing one mutant nitrate reductase (*NR*) gene, could be converted to fully *NR* null mutants (chlorate resistant) following *Agrobacterium*–mediated



transformation. However, none of these null mutants contained T-DNA in the *NR* gene [75]. Mutation of the wild-type *NR* allele must have occurred by some other mechanism.

Another indication that *Agrobacterium* infection may be inherently mutagenic derives from the observation that only ~35% of the T-DNAs in *Arabidopsis* T-DNA insertion libraries co-segregate with a screened mutant phenotype [76,77]. Mutations in the selected lines may be derived from disruptions other than T-DNA insertion into the gene of interest.

Many *Arabidopsis* T-DNA insertion lines contain complex host genome rearrangements that are frequently associated with mis-repair of double-strand DNA breaks. These include inversions, translocations, and other complex rearrangements [78–85]. Similar rearrangements have been detected in transgenic rice [86]. Clark and Krysan [87] noted that approximately 19% of the examined lines from the SALK T-DNA mutant collection contained translocations. The rearrangements of plant genomes following T-DNA integration are reminiscent of the process of chromothripsis resulting from CRISPR–Cas9 mammalian genome editing [88].

### 5. What Is the Mechanism of T-DNA Integration?

Perhaps the most important problem remaining in understanding *Agrobacterium*-mediated transformation is the mechanism of T-DNA integration. As cited above, numerous models of integration have been proposed. What is clear is that homologous recombination is not the mechanism: Despite many kilobases of homology between plant DNA and engineered T-DNA, integration into homologous sequences in the plant genome occurs extremely rarely. This differs from the situation in yeast, where homologous recombination is predominant when homology between T-DNA and the yeast genome is present (see, e.g., [89–91]). Thus, what remains for the T-DNA integration mechanism in plants is some form of non-homologous end-joining (NHEJ) in which T-DNA integration occurs in the absence of large regions of homology, although targeting by microhomology may be used in some circumstances.

Two major NHEJ pathways have been described (e.g., [92–95]). The “classical” (Ku-dependent) pathway utilizes, among other proteins, the Ku70/Ku80 heterodimer to protect the broken DNA ends, and the complex of XRCC4/XLS/DNA ligase IV to repair the break. It is not unusual that, following repair, small deletions, insertions, or nucleotide substitutions occur at or near the break site. Microhomology between the ligated ends is rarely detected. An “alternative” pathway uses microhomology between a region at or near the break site and another sequence (near or distant from the break site) for repair. Participants in this pathway include members of the MRN complex that process broken chromosome ends, the WRN helicase, and a complex of XRCC1 and DNA ligase III (not found in plants) to repair the breaks. DNA polymerase  $\theta$  is a participant in this pathway and is proposed to play a key role in T-DNA integration. Microhomology-mediated end-joining (MMEJ) is frequently referred to as theta-mediated end-joining because of DNA polymerase  $\theta$ 's role in this process. DNA polymerase  $\theta$  has several unusual properties: the protein is made up of both a helicase and a DNA polymerase domain, and the enzyme has a propensity to “template switch”. This latter property allows it to copy DNA from another region of the genome into break sites, generating “filler” DNA sequences in the break. MMEJ is highly mutagenic, frequently generating deletions as sequences flanking DNA break sites search for homologous sequences with which to join. MMEJ also generates chromosomal rearrangements such as inversions and translocations, features commonly associated with T-DNA integration.

Which of these NHEJ pathways, if any, are involved in T-DNA integration? Numerous studies have been published testing stable transformation efficiencies and T-DNA integration characteristics of various *Arabidopsis* and rice NHEJ mutants [6,96–106]. However, with the exception of three publications [102,104,106], all other studies used the frequency of stable transformation as a proxy for T-DNA integration. While detection of stable transformants requires T-DNA integration, it also requires expression of selection marker genes to recover transformed tissue. T-DNA integration may thus occur in the absence of stable

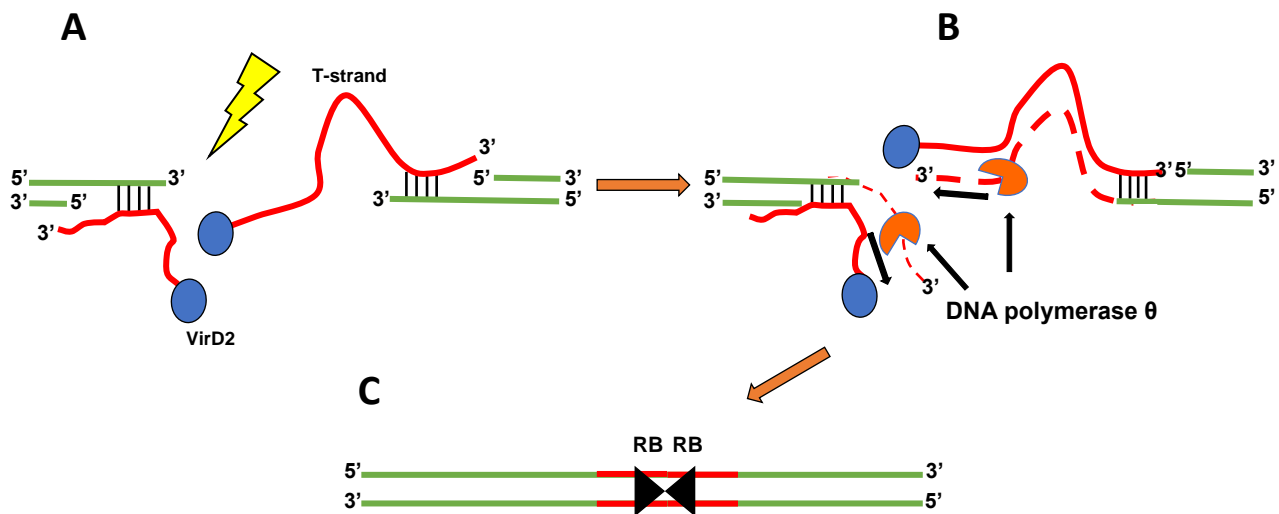
transformation if selection marker genes have been silenced. As noted above, such selection bias can confound experimental interpretations [41,42,107]. An additional complication is that most *Arabidopsis* stable transformation experiments were conducted using a flower-dip protocol. It is well-documented that the importance of *Arabidopsis* genes essential for somatic cell transformation differs from that of germ-line transformation [106,108,109]. Finally, stable transformation efficiencies must be calculated with respect to transient transformation frequencies; a decrease in stable transformation may not indicate that a particular plant mutant has altered stable transformation characteristics if the transient transformation frequency is correspondingly altered. It is particularly important that plant inoculations be conducted with several orders of magnitude different *Agrobacterium* concentrations to avoid a “saturation response” with high bacterial inoculum conditions, thus obscuring differences among wild-type and mutant plant genotypes.

In light of these numerous variables and limitations, it may not be surprising that different laboratories have come to different conclusions with regard to the importance of various plant NHEJ genes for T-DNA integration (or rather, for most studies, stable transformation). Several reports indicate that mutation of the *Arabidopsis* or rice classical (c)NHEJ pathway genes *Ku70*, *Ku80*, or DNA ligase IV (*Lig4*) resulted in lower stable transformation frequencies [6,96,99,100,103]. These studies suggest that these cNHEJ genes are important for T-DNA integration. Other publications indicated that such mutations had little or no effect on stable transformation [97,98]. These studies suggest that these cNHEJ genes are not essential for T-DNA integration. Still other publications, using both *Arabidopsis* and *N. benthamiana*, showed that mutation or down-regulation of several cNHEJ genes, including *Ku70*, *Ku80*, *XRCC4*, and the gene encoding DNA ligase VI (*Lig6*), increased both stable transformation and T-DNA integration into non-selected plant cells [102,104]. These studies suggest that expression of these cNHEJ genes inhibits T-DNA integration, perhaps by speeding the repair of double-strand DNA breaks required for T-DNA integration.

Similarly, individual mutation of two genes associated with MMEJ, *XRCC1* and *PARP2*, did not decrease stable transformation of *Arabidopsis* root tissue ([104]; *PARP1* described in this study has more recently been termed *PARP2*). Mutation of *PARP2* actually increased the frequency of T-DNA integration into the genome of non-selected root cells 2- to 10-fold. The discrepancy between increased T-DNA integration frequency and similar stable transformation frequency of wild-type and *parp2* mutant roots was explained by increased DNA methylation of T-DNA in the *parp2* mutant plants, likely resulting in silencing of the selection genes. This result indicates the importance of investigating T-DNA integration biochemically in non-selected tissue, rather than relying on stable transformation frequency of selected tissue as a proxy for T-DNA integration.

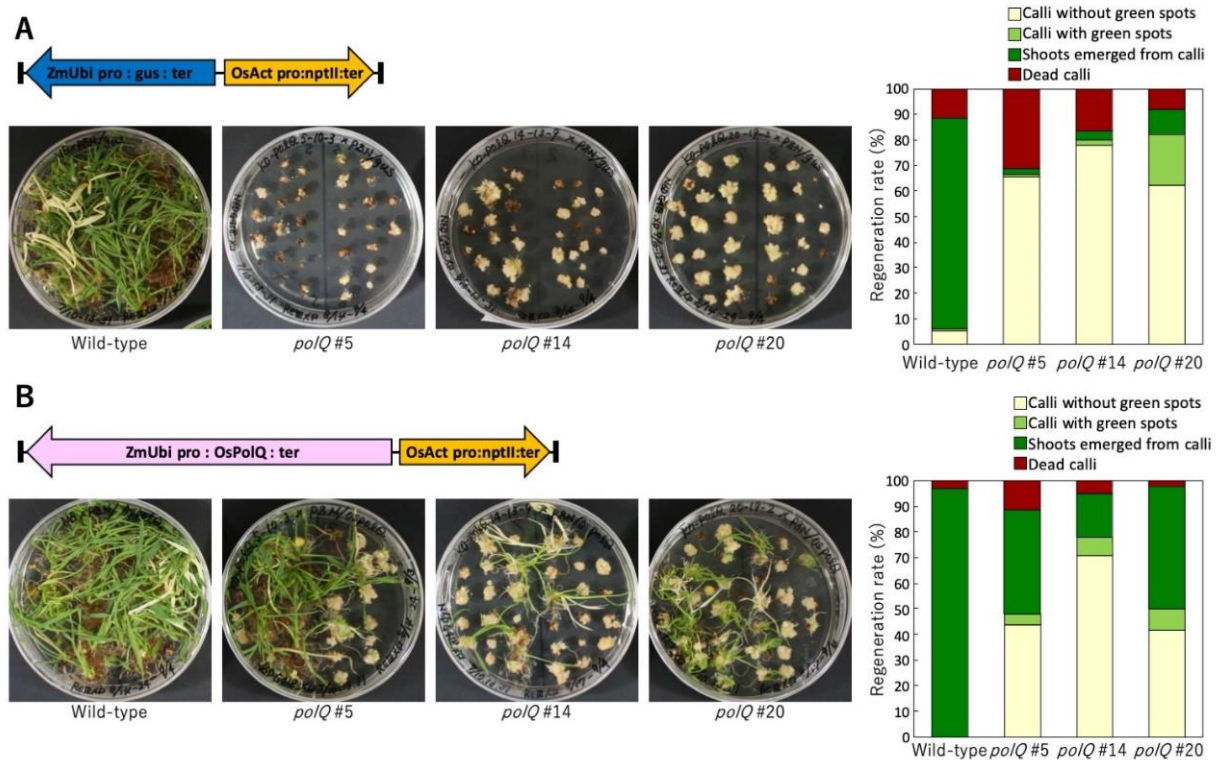
## 6. The Importance of DNA Polymerase $\theta$ for *Agrobacterium*-mediated Transformation and T-DNA Integration

In 2016 van Kregten et al. [7] published a seminal paper in which they proposed an essential function for DNA polymerase  $\theta$  in stable transformation of *Arabidopsis* and T-DNA integration into its genome. These authors examined two DNA polymerase  $\theta$  (*polQ*) mutants, *tebichi (teb) 2* and *teb5*. Although they could not detect differences in transient transformation between wild-type and *polQ* mutant plants, they were not able to obtain any stable transformants of the *polQ* mutants using either a flower-dip transformation protocol or a root transformation protocol requiring selection of transgenic calli and regeneration of plants from these calli. The authors noted that DNA polymerase  $\theta$  can “template switch” during DNA replication, and that it can thereby generate “filler” DNA sequences, a common characteristic of T-DNA/plant DNA junctions at the break site, by copying and joining T-strand DNA and microhomologous plant DNA. They also noted that copying T-strand sequences into both ends of a plant DNA double-strand break could result in integration of T-DNA “head-to-head” (RB-to-RB) dimers, also a common characteristic of many T-DNA insertions (Figure 1). T-DNA integration via theta-mediated end-joining thus became the favored model for T-DNA integration into plant genomes.



**Figure 1.** Model for the involvement of DNA polymerase  $\theta$  in T-DNA integration. (A) Following generation of a double-strand DNA break (lightning bolt), the 5' ends of plant DNA (green lines) are resected. Sequences within single-strand T-DNA molecules (T-strands, red lines), capped with VirD2 protein (blue oval) at their 5' ends, align by microhomology with complementary regions of the exposed single-strand plant DNA. This can occur at both sides of the break (as pictured) or at one side only. (B) DNA polymerase  $\theta$  (orange “pac-man” symbol) copies the T-strand sequence, using the exposed 3' end of the plant DNA as a primer. The newly synthesized T-strand complementary sequence is ligated to the 5' end of the break site, and the 5' end of the T-strand is ligated to the 3' end of the broken plant DNA following VirD2 removal. (C) Final result of integration of two T-DNA molecules in a head-to-head (RB-to-RB) orientation.

Nishizawa-Yokoi et al. [106] re-examined the role of DNA polymerase  $\theta$  in T-DNA integration. Using the same *Arabidopsis* *teb2* and *teb5* mutants used by van Kregten et al. [7], as well as three independent rice *polQ* mutants, this group was able to obtain stable transformants of somatic tissue in all tested *polQ* mutants. Similar to van Kregten et al. [7], they were not able to transform *Arabidopsis* by a flower-dip protocol, except when the incoming T-DNA constitutively expressed a wild-type *PolQ* gene. These authors additionally showed that transient transformation of roots from the *Arabidopsis* *polQ* mutants did decrease relative to transformation of wild-type roots. T-DNA/plant DNA junctions isolated from transformed rice and *Arabidopsis* *polQ* mutant calli had characteristics similar to those isolated from wild-type tissue. Finally, the authors showed that both *Arabidopsis* and rice *polQ* mutants had growth and/or developmental defects; root segments from *Arabidopsis* *polQ* mutants did not form callus well and the calli grew slowly. Calli derived from rice *polQ* mutants did not regenerate plants even under non-transformation and non-selection conditions. The variable penetrance of the *tebichi* phenotype was recently examined and was shown to increase under stress, including replication stress, conditions [110,111]. Similar to the situation with *Arabidopsis* flower-dip transformation, rice *polQ* mutants could be transformed and regenerated into plants if the incoming T-DNA contained a constitutively expressed *PolQ* gene (Figure 2). Thus, transformation and developmental deficiencies resulting from mutation of *PolQ* could be complemented by transient expression of a wild-type *PolQ* gene in both *Arabidopsis* and rice.



**Figure 2.** Regeneration of wild-type and *polQ* mutant rice plants after transformation by an *Agrobacterium* strain lacking or harboring a wild-type rice *PolQ* gene in the T-DNA. Transgenic calli (wild-type Nipponbare or three different *polQ* mutant lines (see [106])) were selected on medium containing 35 mg/L geneticin (G418) and 25 mg/L meropenem for four weeks. The calli were then transferred to regeneration medium (ReIII) lacking geneticin. Left panels, the plates were photographed after further incubation for four weeks. Right panels, the percentage of calli that developed shoots was quantified. (A) Infection by an *Agrobacterium* strain lacking a *PolQ* gene in the T-DNA. (B) Infection by an *Agrobacterium* strain containing a *PolQ* gene in the T-DNA. Data are from Dr. Ayako Nishizawa-Yokoi.

What were the differences between the experiments conducted by these two groups? For *Arabidopsis* transformation, both groups used the same mutant lines, *teb2* and *teb5*. The major difference was the assays used to determine transient and stable transformation. Van Kregten et al. [7] used a stable transformation assay in which transformed *Arabidopsis* root segments were regenerated to green plants under phosphinothricin selection conditions (the T-DNA they used contained a *bar* gene). Nishizawa-Yokoi et al. [106] used an assay in which root segments were regenerated only to the callus stage. Selection consisted of either crown gall tumor formation on phytohormone-free medium (using an oncogenic *Agrobacterium* strain), or callus induction medium containing phosphinothricin (using an *Agrobacterium* strain with a *bar* gene in the T-DNA region). Tumors and calli from the *teb* mutant roots grew more slowly than did tumors/calli from wild-type plants, reflecting the growth phenotype of *polQ Arabidopsis* mutants. However, stable transformants were detected at the frequency of 25–55% that of wild-type plants. Growth of calli, instead of regeneration of selected plants, thus allowed recovery of transformed tissue for further analysis. The average amount of T-DNA integrated into the genome of *teb* mutant calli, that had not been selected for transformation (to avoid selection bias), was 50–90% that of wild-type calli, as determined by droplet digital PCR. Thus, T-DNA integration into the genome of *Arabidopsis polQ* mutant cells was similar to that of wild-type cells. However, it was more difficult to recover tissue from *polQ* mutants than from wild-type tissue to permit molecular analysis.

A second difference between the groups was the assays for transient transformation. Van Kregten et al. [7] did not see any difference between the wild-type and *teb* mutant plants, whereas Nishizawa-Yokoi et al. [106] saw a decrease in transient transformation of

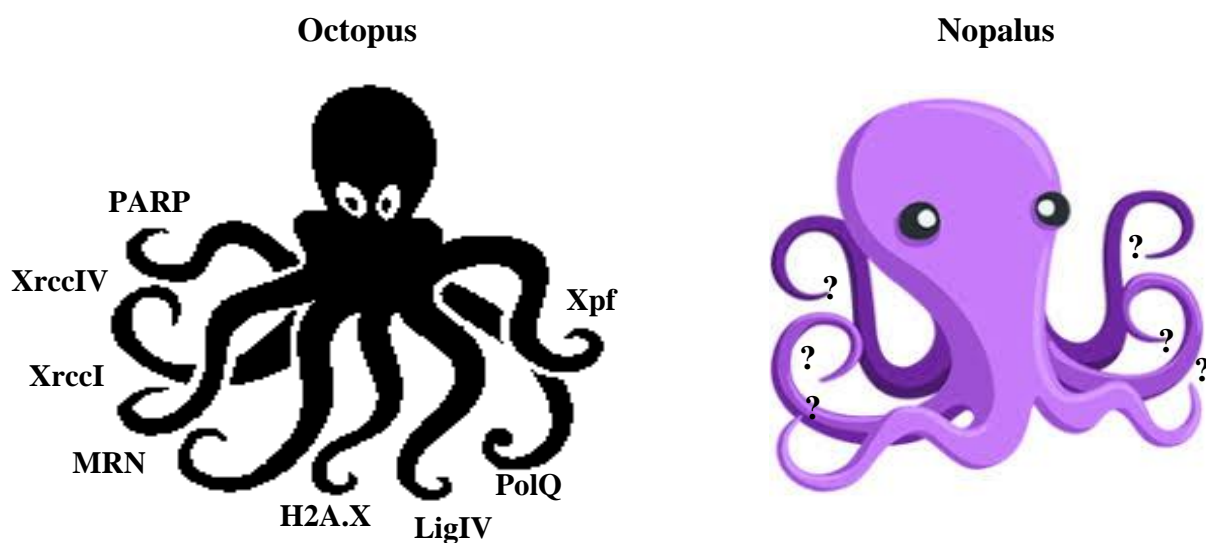


the mutants. Transformation assays can be very sensitive to the bacterial inoculum used; high *Agrobacterium* inoculum concentrations, as used by Van Kregten et al. [7], can mask differences between wild-type and mutant plant responses. Nishizawa-Yokoi et al. [106] used 10-fold serial dilutions of bacteria, starting at a lower bacterial concentration than that used by Van Kregten et al., and were easily able to quantify differences in transient transformation frequencies of wild-type and *teb* mutant plants.

A third difference between the groups was that Nishizawa-Yokoi et al. [106] additionally showed that mutation of *polQ* in rice also allowed stable transformation, although at a reduced frequency to that of wild-type rice tissue.

### 7. T-DNA Integration: An Octopus or a Nopalus?

One could consider at least two explanations for the discrepancies in the literature regarding the importance of individual DNA repair genes and pathways for T-DNA integration: (1) multiple DNA repair pathways could be used for integration. No one pathway is essential, but many could contribute. Thus, like the multiarmed octopus, T-DNA may “grab” any pathway it can find to integrate. This model would predict that disruption of any one pathway, or even simultaneous disruption of multiple pathways, would perhaps decrease but not eliminate T-DNA integration, depending on the importance of that pathway for integration; or (2) none of the known DNA repair pathways play a role in T-DNA integration. Like the mythical nopalus, some as yet unknown pathway mediates integration (Figure 3).



**Figure 3.** “Fantastic” models for mechanisms of T-DNA integration. Left, the octopus which grabs proteins from any NHEJ pathway it can to facilitate T-DNA integration. Right, the mythical nopalus, representing unknown pathway(s) for T-DNA integration. (The author thanks Barbara Hohn for suggesting the term “octopus” to describe one T-DNA integration model, and John Kemp for suggesting, in 1979, the term “nopalus” as a sarcastic alternative to “octopus”, from which octopine is derived, for the origin of nopaline. “Nopal” is a Spanish name for *Opuntia* cacti. Nopaline was first isolated from *Opuntia* crown gall tumors).

What are the data to suggest that multiple DNA repair pathways may contribute to T-DNA integration? Several groups have reported the effect of simultaneous mutation of multiple DNA repair genes, functioning in different repair pathways, on stable transformation or T-DNA integration. Simultaneous mutation of *Arabidopsis ku80* (important for classical NHEJ) and *parp2* (important for alternative NHEJ) had little effect on either stable transformation of *Arabidopsis* root segments or T-DNA integration into the *Arabidopsis* genome [106]. Mestiri et al. [103] examined transformation of various combinations of *Arabidopsis* DNA repair mutants, including *ku80* (important for the classical NHEJ pathway), *xrcc1* (important for the alternative or MMEJ pathway), *xpf* (important for NHEJ and

homologous recombination (HR)), and *xrcc2* (important for HR). Increasing the number of mutant repair/recombination genes (from *ku80*, to *ku80/xrcc1*, to *ku80/xrcc1/xpf*, to *ku80/xrcc1/xpf/xrcc2*) progressively decreased stable transformation from 2.6- to 5.4-fold, with little decrease in transient transformation. However, stable transformation was not eliminated, indicating that some other pathway(s) must exist for stable transformation to occur.

More recently, our laboratory has re-examined these same mutants (single, double, triple, and quadruple; [103]) for transient and stable transformation, for T-DNA integration under non-selective growth conditions, and for the growth and developmental phenotypes of these mutants [112]. As more mutations in DNA repair pathways were introduced, the plants became progressively more sensitive to light, and mutant plants grown under moderate light intensity displayed growth and developmental phenotypes. Using multiple serial dilutions of *Agrobacterium* for inoculation, major decreases in transient transformation, as well as stable transformation, were revealed in these mutants. However, T-DNA/plant DNA junctions isolated from single and *ku80/xrcc1* double mutants had characteristics similar to those isolated from transformed wild-type Col-0 plants. Finally, using droplet digital PCR to evaluate *Agrobacterium*-infected calli grown under non-selective conditions, little difference existed in the amount of T-DNA integrated into the genomes of wild-type calli and calli derived from the various DNA repair mutants. Although these results differ quantitatively from those of Mestiri et al. [103], they confirm that simultaneous mutation of genes in these various DNA repair and recombination pathways do not eliminate either stable transformation or T-DNA integration. Therefore, there must be some other pathway available for integration to occur.

## 8. Where Do We Go from Here?

Studies on the importance of DNA repair genes and pathways in T-DNA integration have, for the most part, depended on a genetic approach to examine stable *Agrobacterium*-mediated transformation. The potential problems of using stable transformation as a proxy for T-DNA integration have been discussed above. However, a genetic approach may also have limitations: Genes important for T-DNA integration may be essential for cell viability, and homozygous mutation of these genes may be lethal to the organism. An example of a gene that may fit this category is that encoding DNA ligase I (*Lig1*). Because of its role in DNA replication, homozygous mutants in *lig1* are lethal [113]. Therefore, one may not expect to recover a homozygous *lig1* mutant when screening for T-DNA integration-deficient plants. It may be possible to screen partial loss-of-function *lig1* mutants for disruption of T-DNA integration. However, considering the important role Lig1 protein plays in fundamental cellular processes, interpretation of these integration data may be difficult.

A complementary approach to understanding proteins and pathways essential for T-DNA integration may be the use of biochemistry to dissect the integration process. Ultimately, development of an in vitro T-DNA integration assay, with fully characterized components, may be required to dissect the integration process. The beginning of such an assay was described previously [26]. However, this assay used a mostly uncharacterized plant cell lysate.

A major difficulty in studying T-DNA integration is that it is random, and therefore difficult to analyze biochemically in a large plant genome. Because double-strand DNA breaks tend to “trap” T-DNA, one approach to investigating the integration process may be to use CRISPR technology to generate a double-strand DNA break at a precise genomic location just prior to/concomitant with transformation. This will allow scientists to concentrate their biochemical methodologies on a specific locus. For example, one could analyze *Agrobacterium* and/or plant proteins, including those affiliated with T-strands, that are attracted to these breaks.

Another approach would be to use specific and more readily accessible molecular forms of T-DNA to investigate the roles of proteins (both known or surmised) in alternative

“integration-like” mechanisms that accurately reflect T-DNA integration. To this end, the characterization of T-circles isolated from *Agrobacterium*-infected plants may prove fruitful. T-circles are double-strand circular molecules that contain T-DNA that is joined at or near T-DNA RB and LB sequences [114,115]. They are formed in the plant and have many characteristics of T-DNA/plant DNA junctions, although they are not integrated into the plant genome. For example, T-circles can consist of T-DNA dimers in direct (RB-to-LB) or inverted (RB-to-RB or LB-to-LB) orientation, similar to what is frequently seen with integrated T-DNAs. RB sequences can be precise (i.e., cut between nucleotides 3 and 4 of the 25 bp RB) or have small deletions. Deletions at the LB are more common and more extensive. Filler DNA can appear between the RB and LB sequences [115].

More recent experiments analyzing hundreds of T-circles confirmed and extended information about T-circle RB and LB junctions [116]. These additional features include identification of filler DNA between the RB and LB as coming from the plant genome or from other replicons in *Agrobacterium*, including the pAtC58 “cryptic” plasmid, the vir helper plasmid, or regions from the binary vector backbone. Filler DNA from these genomic sources was detected previously (Supplementary Table S1 of [106]). Some of these vector backbone sequences are in “inverted orientation” relative to the orientation of the border sequences, a feature found in integrated T-DNAs [44]. Microhomology frequently exists between this filler DNA and the adjacent T-DNA sequences. In addition, T-circles can be recovered from *Arabidopsis ku80* mutant plants at the same frequency that they are recovered from wild-type plants, and the RB–LB T-circle junctions are similar regardless of whether they are isolated from wild-type or *ku80* mutant plants. These characteristics are the same as those seen in integrated T-DNA/plant DNA junctions [104,112]. Thus, in all aspects examined to date, T-circle border junctions resemble those of integrated T-DNA. Because of these similarities, T-circles may serve as a more readily accessible proxy for studying the mechanism of, and the proteins important for, T-DNA integration.

### 9. Important Remaining Questions to Answer Regarding T-DNA Integration

Many important questions remain for investigating T-DNA integration and the role that plant DNA break repair processes may play. Among these are:

1. Are plant single-strand nicks, double-strand breaks, neither, or both required for T-DNA integration?
2. Are there particular aspects of plant chromatin that are more conducive for integration to occur?
3. What are the roles, if any, of *Agrobacterium* virulence effector proteins (such as VirD2, VirE2, and possibly others) in T-DNA integration?
4. What plant proteins are important for integration? Related to this question, what (if any) plant DNA repair pathway(s) is/are used for integration?
5. Transient transformation of most plant species and tissues is considerably more efficient than is stable transformation. Why is this? If many T-strands enter the nucleus and can initially be converted to double-strand transcription-competent forms (either linear or circular), why is T-DNA integration relatively rare in these nuclei? Could greater expression of plant DNA repair genes important for T-DNA integration increase the percentage of stably transformed cells?
6. What form of T-DNA (single-strand, double-strand linear, double-strand circular) is the substrate or template for integration?
7. What is the molecular basis for differences in the frequency of T-DNA integration among plant species, or even among varieties/cultivars of the same species?
8. During transformation, many plant cells are exposed to *Agrobacterium*, but only a few may be transformed (either transiently or stably). What is the basis for plant cell transformation competency?

**Funding:** Research in the author's laboratory is funded by the US National Science Foundation (award numbers 1543888, 1848434, and 2006668).

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Mayerhofer, R.; Koncz-Kalman, Z.; Nawrath, C.; Bakkeren, G.; Cramer, A.; Angelis, K.; Redei, G.P.; Schell, J.; Hohn, B.; Koncz, C. T-DNA integration: A mode of illegitimate recombination in plants. *EMBO J.* **1991**, *10*, 697–704. [[CrossRef](#)]
2. Tinland, B.; Hohn, B. Recombination between prokaryotic and eukaryotic DNA: Integration of *Agrobacterium tumefaciens* T-DNA into the plant genome. In *Genetic Engineering*; Setlow, J.K., Ed.; Plenum Press: New York, NY, USA, 1995; pp. 209–229.
3. Tinland, B. The integration of T-DNA into plant genomes. *Trends Plant Sci.* **1996**, *1*, 178–184. [[CrossRef](#)]
4. Tzfira, T.; Li, J.; Lacroix, B.; Citovsky, V. *Agrobacterium* T-DNA integration: Molecules and models. *Trends Genet.* **2004**, *20*, 375–383. [[CrossRef](#)]
5. Bilichak, A.; Yao, Y.; Kovalchuk, I. Transient down-regulation of the RNA silencing machinery increases efficiency of *Agrobacterium*-mediated transformation of *Arabidopsis*. *Plant Biotechnol. J.* **2014**, *12*, 590–600. [[CrossRef](#)] [[PubMed](#)]
6. Saika, H.; Nishizawa-Yokoi, A.; Toki, S. The non-homologous end-joining pathway is involved in stable transformation in rice. *Front. Plant Sci.* **2014**, *5*, 560. [[CrossRef](#)]
7. van Kregten, M.; de Pater, S.; Romeijn, R.; van Schendel, R.; Hooykaas, P.J.J.; Tijsterman, M. T-DNA integration in plants results from polymerase-theta-mediated DNA repair. *Nature Plants* **2016**, *2*, 16164. [[CrossRef](#)] [[PubMed](#)]
8. Gelvin, S.B. Integration of *Agrobacterium* T-DNA into the plant genome. *Annu. Rev. Genet.* **2017**, *51*, 195–217. [[CrossRef](#)]
9. Lacroix, B.; Citovsky, V. Pathways of DNA transfer to plants from *Agrobacterium tumefaciens* and related bacterial species. *Annu. Rev. Phytopathol.* **2019**, *57*, 231–251. [[CrossRef](#)]
10. Dong, O.X.; Ronald, P.C. Targeted DNA insertion in plants. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2004834117. [[CrossRef](#)] [[PubMed](#)]
11. Yanofsky, M.F.; Porter, S.G.; Young, C.; Albright, L.M.; Gordon, M.P.; Nester, E.W. The *virD* operon of *Agrobacterium tumefaciens* encodes a site-specific endonuclease. *Cell* **1986**, *47*, 471–477. [[CrossRef](#)]
12. Stachel, S.E.; Timmerman, B.; Zambryski, P. Generation of single-stranded T-DNA molecules during the initial stages of T-DNA transfer from *Agrobacterium tumefaciens* to plant cells. *Nature* **1986**, *322*, 706–712. [[CrossRef](#)]
13. Jayaswal, R.K.; Veluthambi, K.; Gelvin, S.B.; Slightom, J.L. Double-stranded cleavage of T-DNA and generation of single-stranded T-DNA molecules in *Escherichia coli* by a *virD*-encoded border-specific endonuclease from *Agrobacterium tumefaciens*. *J. Bacteriol.* **1987**, *169*, 5035–5045. [[CrossRef](#)]
14. Veluthambi, K.; Jayaswal, R.K.; Gelvin, S.B. Virulence genes A, G, and D mediate the double-stranded border cleavage of T-DNA from the *Agrobacterium* Ti plasmid. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 1881–1885. [[CrossRef](#)]
15. Veluthambi, K.; Ream, W.; Gelvin, S.B. Virulence genes, borders, and overdrive generate single-stranded T-DNA molecules from the A6 Ti plasmid of *Agrobacterium tumefaciens*. *J. Bacteriol.* **1988**, *170*, 1523–1532. [[CrossRef](#)] [[PubMed](#)]
16. Yadav, N.S.; Van der Leyden, J.; Bennett, D.R.; Barnes, W.M.; Chilton, M.-D. Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 6322–6326. [[CrossRef](#)] [[PubMed](#)]
17. Wang, K.; Herrera-Estrella, L.; Van Montagu, M.; Zambryski, P. Right 25 bp terminus sequence of the nopaline T-DNA is essential for and determines direction of DNA transfer from *Agrobacterium* to the plant genome. *Cell* **1984**, *38*, 455–462. [[CrossRef](#)]
18. Herrera-Estrella, A.; Chen, Z.-M.; Van Montagu, M.; Wang, K. VirD proteins of *Agrobacterium tumefaciens* are required for the formation of a covalent DNA-protein complex at the 5' terminus of T-strand molecules. *EMBO J.* **1988**, *7*, 4055–4062. [[CrossRef](#)] [[PubMed](#)]
19. Ward, E.R.; Barnes, W.M. VirD2 protein of *Agrobacterium tumefaciens* very tightly linked to the 5' end of T-strand DNA. *Science* **1988**, *242*, 927–930. [[CrossRef](#)]
20. Young, C.; Nester, E.W. Association of the VirD2 protein with the 5' end of T strands in *Agrobacterium tumefaciens*. *J. Bacteriol.* **1988**, *170*, 3367–3374. [[CrossRef](#)] [[PubMed](#)]
21. Durrenberger, F.; Cramer, A.; Hohn, B.; Koukolikova-Nicola, Z. Covalently bound VirD2 protein of *Agrobacterium tumefaciens* protects the T-DNA from exonucleolytic degradation. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 9154–9158. [[CrossRef](#)] [[PubMed](#)]
22. Howard, E.A.; Winsor, B.A.; De Vos, G.; Zambryski, P. Activation of the T-DNA transfer process in *Agrobacterium* results in the generation of a T-strand-protein complex: Tight association of VirD2 with the 5' ends of T-strands. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 4017–4021. [[CrossRef](#)]
23. Tinland, B.; Hohn, B.; Puchta, H. *Agrobacterium tumefaciens* transfers single-stranded transferred DNA (T-DNA) into the plant cell nucleus. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 8000–8004. [[CrossRef](#)]
24. Yusibov, V.M.; Steck, T.R.; Gupta, V.; Gelvin, S.B. Association of single-stranded transferred DNA from *Agrobacterium tumefaciens* with tobacco cells. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 2994–2998. [[CrossRef](#)] [[PubMed](#)]
25. Pansegrau, W.; Schoumacher, F.; Hohn, B.; Lanka, E. Site-specific cleavage and joining of single-stranded DNA by VirD2 protein of *Agrobacterium tumefaciens* Ti plasmids: Analogy to bacterial conjugation. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 11538–11542. [[CrossRef](#)]
26. Ziemienowicz, A.; Tinland, B.; Bryant, J.; Gloeckler, V.; Hohn, B. Plant enzymes but not *Agrobacterium* VirD2 mediate T-DNA ligation in vitro. *Mol. Cell. Biol.* **2000**, *20*, 6317–6322. [[CrossRef](#)] [[PubMed](#)]



27. Shurvinton, C.E.; Hodges, L.; Ream, W. A nuclear localization signal and the C-terminal omega sequence in the *Agrobacterium tumefaciens* VirD2 endonuclease are important for tumor formation. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 11837–11841. [[CrossRef](#)]
28. Narasimhulu, S.B.; Deng, X.-B.; Sarria, R.; Gelvin, S.B. Early transcription of *Agrobacterium* T-DNA genes in tobacco and maize. *Plant Cell* **1996**, *8*, 873–886. [[PubMed](#)]
29. Mysore, K.S.; Bassuner, B.; Deng, X.-b.; Darbinian, N.S.; Motchoulski, A.; Ream, W.; Gelvin, S.B. Role of the *Agrobacterium tumefaciens* VirD2 protein in T-DNA transfer and integration. *Mol. Plant.-Microbe Interact.* **1998**, *11*, 668–683. [[CrossRef](#)] [[PubMed](#)]
30. Tinland, B.; Schoumacher, F.; Gloeckler, V.; Bravo-Angel, A.M.; Hohn, B. The *Agrobacterium tumefaciens* virulence D2 protein is responsible for precise integration of T-DNA into the plant genome. *EMBO J.* **1995**, *14*, 3585–3595. [[CrossRef](#)]
31. Bravo-Angel, A.M.; Hohn, B.; Tinland, B. The omega sequence of VirD2 is important but not essential for efficient transfer of T-DNA by *Agrobacterium tumefaciens*. *Mol. Plant.-Microbe Interact.* **1997**, *11*, 57–63. [[CrossRef](#)]
32. Nuti, M.P.; Ledebuer, A.M.; Durante, M.; Nuti-Ronchi, V.; Schilperoort, R.A. Detection of Ti-plasmid sequences in infected tissues by in situ hybridization. *Plant Sci. Lett.* **1980**, *18*, 1–6. [[CrossRef](#)]
33. Ambros, P.F.; Matzke, A.J.M.; Matzke, M.A. Localization of *Agrobacterium rhizogenes* T-DNA in plant chromosomes by in situ hybridization. *EMBO J.* **1986**, *5*, 2073–2078. [[CrossRef](#)]
34. Wang, J.; Lewis, M.E.; Whallon, J.H.; Sink, K.C. Chromosomal mapping of T-DNA inserts in transgenic *Petunia* by in situ hybridization. *Transgen. Res.* **1995**, *4*, 241–246. [[CrossRef](#)]
35. Koncz, C.; Martini, N.; Mayerhofer, R.; Koncz-Kalman, Z.; Korber, H.; Redei, G.P.; Schell, J. High-frequency T-DNA-mediated gene tagging in plants. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 8467–8471. [[CrossRef](#)]
36. Brunaud, V.; Balzergue, S.; Dubreucq, B.; Aubourg, S.; Samson, F.; Chauvin, S.; Bechtold, N.; Cruaud, C.; DeRose, R.; Pelletier, G.; et al. T-DNA integration into the *Arabidopsis* genome depends on sequences of pre-insertion sites. *EMBO Rep.* **2002**, *3*, 1152–1157. [[CrossRef](#)] [[PubMed](#)]
37. Szabados, L.; Kovacs, I.; Oberschall, A.; Abraham, E.; Kerekes, I.; Zsigmond, L.; Nagy, R.; Alvarado, M.; Krasovskaja, I.; Gal, M.; et al. Distribution of 1000 sequenced T-DNA tags in the *Arabidopsis* genome. *Plant J.* **2002**, *32*, 233–242. [[CrossRef](#)]
38. Chen, S.; Jin, W.; Wang, M.; Zhang, F.; Zhou, J.; Jia, Q.; Wu, Y.; Liu, F.; Wu, P. Distribution and characterization of over 1000 T-DNA tags in rice genome. *Plant J.* **2003**, *36*, 105–113. [[CrossRef](#)]
39. Sallaud, C.; Gay, C.; Larmande, P.; Bes, M.; Piffanelli, P.; Piegu, B.; Droc, G.; Regad, F.; Bourgeois, E.; Meynard, D.; et al. High throughput T-DNA insertion mutagenesis in rice: A first step towards in silico reverse genetics. *Plant J.* **2004**, *39*, 450–464. [[CrossRef](#)]
40. Schneeberger, R.G.; Zhang, K.; Tatarinova, T.; Troukhan, M.; Kwok, S.F.; Drais, J.; Klinger, K.; Orejudos, F.; Macy, K.; Bhakta, A.; et al. *Agrobacterium* T-DNA integration in *Arabidopsis* is correlated with DNA sequence compositions that occur frequently in gene promoter regions. *Funct. Integr. Genom.* **2005**, *5*, 240–253. [[CrossRef](#)]
41. Kim, S.I.; Veena; Gelvin, S.B. Genome-wide analysis of *Agrobacterium* T-DNA integration sites in the *Arabidopsis* genome generated under non-selective conditions. *Plant J.* **2007**, *51*, 779–791. [[CrossRef](#)]
42. Shilo, S.; Tripathi, P.; Melamed-Bessudo, C.; Tzfadia, O.; Muth, T.R.; Levy, A.A. T-DNA-genome junctions form early after infection and are influenced by the chromatin state of the host genome. *PLoS Genet.* **2017**, *13*, e1006875. [[CrossRef](#)]
43. Windels, P.; De Buck, S.; Van Bockstaele, E.; De Loose, M.; Depicker, A. T-DNA integration in *Arabidopsis* chromosomes: Presence and origin of filler DNA sequences. *Plant Physiol.* **2003**, *133*, 2061–2068. [[CrossRef](#)]
44. Muller, A.E.; Atkinson, R.G.; Sandoval, R.B.; Jorgensen, R.A. Microhomologies between T-DNA ends and target sites often occur in inverted orientation and may be responsible for the high frequency of T-DNA-associated inversions. *Plant Cell Rep.* **2007**, *26*, 617–630. [[CrossRef](#)] [[PubMed](#)]
45. Kleinboelting, N.; Huep, G.; Appelhagen, I.; Viehoveer, P.; Li, Y.; Weisshaar, B. The structural features of thousands of T-DNA insertion sites are consistent with a double-strand break repair based insertion mechanism. *Mol. Plant* **2015**, *8*, 1651–1664. [[CrossRef](#)]
46. Nam, J.; Mysore, K.S.; Zheng, C.; Knue, M.K.; Matthyse, A.G.; Gelvin, S.B. Identification of T-DNA tagged *Arabidopsis* mutants that are resistant to transformation by *Agrobacterium*. *Mol. Gen. Genet.* **1999**, *261*, 429–438. [[CrossRef](#)] [[PubMed](#)]
47. Mysore, K.S.; Nam, J.; Gelvin, S.B. An *Arabidopsis* histone H2A mutant is deficient in *Agrobacterium* T-DNA integration. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 948–953. [[CrossRef](#)] [[PubMed](#)]
48. Zheng, Y.; He, X.-W.; Ying, Y.-H.; Lu, J.-F.; Gelvin, S.B.; Shou, H.-X. Expression of the *Arabidopsis thaliana* histone gene *AtHTA1* enhances rice transformation efficiency. *Mol. Plant* **2009**, *2*, 832–837. [[CrossRef](#)] [[PubMed](#)]
49. Yi, H.; Mysore, K.S.; Gelvin, S. Expression of the *Arabidopsis* histone H2A-1 gene correlates with susceptibility to *Agrobacterium* transformation. *Plant J.* **2002**, *32*, 285–298. [[CrossRef](#)]
50. Yi, H.; Sardesai, N.; Fujinuma, T.; Chan, C.W.; Veena; Gelvin, S.B. Constitutive expression exposes functional redundancy between the *Arabidopsis* histone H2A gene *HTA1* and other H2A gene family members. *Plant Cell* **2006**, *18*, 1575–1589. [[CrossRef](#)]
51. Tenea, G.N.; Spantzel, J.; Lee, L.-Y.; Zhu, Y.; Lin, K.; Johnson, S.J.; Gelvin, S.B. Overexpression of several *Arabidopsis* histone genes increases *Agrobacterium*-mediated transformation and transgene expression in plants. *Plant Cell* **2009**, *21*, 3350–3367. [[CrossRef](#)]
52. Gelvin, S.B.; Kim, S.-I. Effect of chromatin upon *Agrobacterium* T-DNA integration and transgene expression. *Biophys. Biochem. Acta* **2007**, *1769*, 410–421. [[CrossRef](#)] [[PubMed](#)]
53. Wolterink-van Loo, S.; Escamilla Ayala, A.A.; Hooykaas, P.J.J.; van Heudsdn, G.P.H. Interaction of the *Agrobacterium tumefaciens* virulence protein VirD2 with histones. *Microbiology* **2015**, *161*, 401–410. [[CrossRef](#)] [[PubMed](#)]

54. Howard, E.; Citovsky, V. The emerging structure of the *Agrobacterium* T-DNA transfer complex. *BioEssays* **1990**, *12*, 103–108. [[CrossRef](#)]
55. Tzfira, T.; Vaidya, M.; Citovsky, V. VIP1, an *Arabidopsis* protein that interacts with *Agrobacterium* VirE2, is involved in VirE2 nuclear import and *Agrobacterium* infectivity. *EMBO J.* **2001**, *20*, 3596–3607. [[CrossRef](#)]
56. Li, J.; Krichevsky, A.; Vaidya, M.; Tzfira, T.; Citovsky, V. Uncoupling of the functions of the *Arabidopsis* VIP1 protein in transient and stable plant genetic transformation by *Agrobacterium*. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5733–5738. [[CrossRef](#)]
57. Loyter, A.; Rosenbluh, J.; Zakai, N.; Li, J.; Kozlovsky, S.V.; Tzfira, T.; Citovsky, V. The plant VirE2 interacting protein 1. A molecular link between the *Agrobacterium* T-complex and the host cell chromatin? *Plant Physiol.* **2005**, *138*, 1318–1321. [[CrossRef](#)]
58. Lacriox, B.; Loyter, A.; Citovsky, V. Association of the *Agrobacterium* T-DNA-protein complex with plant nucleosomes. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 15429–15434. [[CrossRef](#)]
59. Shi, Y.; Lee, L.-Y.; Gelvin, S.B. Is VIP1 important for *Agrobacterium*-mediated transformation? *Plant J.* **2014**, *79*, 848–860. [[CrossRef](#)]
60. Lapham, R.; Lee, L.-Y.; Tsugama, D.; Lee, S.; Mengiste, T.; Gelvin, S.B. VIP1 and its homologs are not required for *Agrobacterium*-mediated transformation, but play a role in *Botrytis* and salt stress responses. *Front. Plant Sci.* **2018**, *9*, 749. [[CrossRef](#)]
61. Crane, Y.M.; Gelvin, S.B. RNAi-mediated gene silencing reveals involvement of *Arabidopsis* chromatin-related genes in *Agrobacterium*-mediated root transformation. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 15156–15161. [[CrossRef](#)]
62. Soltani, J.; van Heudsdonck, G.P.H.; Hooykaas, P.J.J. Deletion of host histone acetyltransferases and deacetylases strongly affects *Agrobacterium*-mediated transformation of *Saccharomyces cerevisiae*. *FEMS Microbiol. Lett.* **2009**, *298*, 228–233. [[CrossRef](#)]
63. Köhler, F.; Cardon, G.; Pöhlman, M.; Gill, R.; Schieder, O. Enhancement of transformation rates in higher plants by low-dose irradiation: Are DNA repair systems involved in the incorporation of exogenous DNA into the plant genome? *Plant Mol. Biol.* **1989**, *12*, 189–199. [[CrossRef](#)]
64. Gorbunova, V.; Levy, A.A. Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. *Nucl. Acids Res.* **1997**, *25*, 4650–4657. [[CrossRef](#)] [[PubMed](#)]
65. Shirley, B.W.; Hanley, S.; Goodman, H.M. Effects of ionizing radiation on a plant genome: Analysis of two *Arabidopsis* transparent testa mutations. *Plant Cell* **1992**, *4*, 333–347. [[PubMed](#)]
66. Somers, D.A.; Makarevitch, I. Transgene integration in plants: Poking or patching holes in promiscuous genomes? *Curr. Opin. Biotechnol.* **2004**, *15*, 126–131. [[CrossRef](#)] [[PubMed](#)]
67. Salomon, S.; Puchta, H. Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *EMBO J.* **1998**, *17*, 6086–6095. [[CrossRef](#)]
68. Chilton, M.-D.; Que, Q. Targeted integration of T-DNA into the tobacco genome at double-stranded breaks: New insights on the mechanism of T-DNA integration. *Plant Physiol.* **2003**, *133*, 956–965. [[CrossRef](#)]
69. Tzfira, T.; Frankman, L.R.; Vaidya, M.; Citovsky, V. Site-specific integration of *Agrobacterium tumefaciens* T-DNA via double-stranded intermediates. *Plant Physiol.* **2003**, *133*, 1011–1023. [[CrossRef](#)]
70. Zhang, Q.; Xing, H.-L.; Wang, Z.-P.; Zhang, H.-Y.; Yang, F.; Wang, X.-C.; Chen, Q.-J. Potential high-frequency off-target mutagenesis induced by CRISPR/Cas9 in *Arabidopsis* and its prevention. *Plant Mol. Biol.* **2018**, *96*, 445–456. [[CrossRef](#)]
71. Lee, K.; Eggenberger, A.L.; Banakar, R.; McCaw, M.; Zhu, H.; Main, M.; Kang, M.; Gelvin, S.B.; Wang, K. CRISPR/Cas9 targeted T-DNA integration in rice. *Plant Mol. Biol.* **2019**, *99*, 317–328. [[CrossRef](#)]
72. Song, J.; Bent, A.F. Microbial pathogens trigger host DNA double-strand breaks whose abundance is reduced by plant defense responses. *PLoS Pathog.* **2014**, *10*, e1004030. [[CrossRef](#)]
73. Joseph, J.T.; Chandhini, S.; Das, S.; Mysore, K.S.; Shah, J.M. Methylation status of *Arabidopsis* DNA repair gene promoters during *Agrobacterium* infection reveals epigenetic changes in three generations. *Plant Mol. Biol. Rep.* **2021**. [[CrossRef](#)]
74. Wei, F.-J.; Lee, L.-Y.; Hsing, Y.-L.; Gelvin, S.B.; Academia Sinica, Taipei, Taiwan and Purdue University, West Lafayette, IN, USA. Unpublished work, 2021.
75. Marton, L.; Hroudá, M.; Pecsvaradi, A.; Czako, M. T-DNA-independent mutations induced in transformed plant cells during *Agrobacterium* co-cultivation. *Transgen. Res.* **1994**, *3*, 317–325. [[CrossRef](#)]
76. Budziszewski, G.J.; Lewis, S.P.; Glover, L.W.; Reineke, J.; Jones, G.; Ziemnik, L.S.; Lonowski, J.; Nyfeler, B.; Aux, G.; Zhou, Q.; et al. *Arabidopsis* genes essential for seedling viability: Isolation of insertional mutants and molecular cloning. *Genetics* **2001**, *159*, 1765–1778. [[CrossRef](#)]
77. McElver, J.; Tzafrir, I.; Aux, G.; Rogers, R.; Ashby, C.; Smith, K.; Thomas, C.; Schetter, A.; Zhou, Q.; Cushman, M.A.; et al. Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana*. *Genetics* **2001**, *159*, 1751–1763. [[CrossRef](#)]
78. Castle, L.A.; Errampalli, D.; Atherton, T.L.; Franzmann, L.H.; Yoon, E.S.; Meinke, D.W. Genetic and molecular characterization of embryonic mutants identified following seed transformation in *Arabidopsis*. *Mol. Gen. Genet.* **1993**, *241*, 504–514. [[CrossRef](#)] [[PubMed](#)]
79. Nacry, P.; Camilleri, C.; Courtial, B.; Caboche, M.; Bouchez, D. Major chromosomal rearrangements induced by T-DNA transformation in *Arabidopsis*. *Genetics* **1998**, *149*, 641–650. [[CrossRef](#)]
80. Tax, F.E.; Vernon, D.M. T-DNA-associated duplication/translocations in *Arabidopsis*. Implications for mutant analysis and functional genomics. *Plant Physiol.* **2001**, *126*, 1527–1538. [[CrossRef](#)]
81. Lafleur, J.; Degroote, F.; Depeiges, A.; Picard, G. A reciprocal translocation, induced by a canonical integration of a single T-DNA, interrupts the HMG-I/Y *Arabidopsis thaliana* gene. *Plant Physiol. Biochem.* **2004**, *42*, 171–179. [[CrossRef](#)] [[PubMed](#)]

82. Curtis, M.J.; Belcram, K.; Bollmann, S.R.; Tominey, C.M.; Hoffman, P.D.; Mercier, R.; Hays, J.B. Reciprocal chromosome translocation associated with T-DNA-insertion mutation in *Arabidopsis*: Genetic and cytological analyses of consequences for gametophyte development and for construction of doubly mutant lines. *Planta* **2009**, *229*, 731–745. [[CrossRef](#)]
83. Ruprecht, C.; Carroll, A.; Persson, S. T-DNA-induced chromosomal translocation in *feronia* and *anxur2* mutants reveal implications for the mechanism of collapsed pollen due to chromosomal rearrangements. *Mol. Plant* **2014**, *7*, 1591–1594. [[CrossRef](#)]
84. Hu, Y.; Chen, Z.; Zhung, C.; Huang, J. Cascade of chromosomal rearrangements caused by a heterogeneous T-DNA integration supports the double-strand break repair model for T-DNA integration. *Plant J.* **2017**, *90*, 954–965. [[CrossRef](#)] [[PubMed](#)]
85. Jupe, F.; Rivkin, A.C.; Michael, T.P.; Zander, M.; Motley, S.T.; Sandoval, J.P.; Slotkin, R.K.; Chien, H.; Castanon, R.; Nery, J.R.; et al. The complex architecture and epigenomic impact of plant T-DNA insertions. *PLoS Genet.* **2019**, *15*, e1007819. [[CrossRef](#)] [[PubMed](#)]
86. Majhi, B.B.; Shah, J.M.; Veluthambi, K. A novel T-DNA integration in rice involving two interchromosomal translocations. *Plant Cell Rep.* **2014**, *33*, 929–944. [[CrossRef](#)]
87. Clark, D.A.; Krysan, P.J. Chromosomal translocations are a common phenomenon in *Arabidopsis thaliana* T-DNA insertion lines. *Plant J.* **2010**, *64*, 990–1001. [[CrossRef](#)] [[PubMed](#)]
88. Leibowitz, M.L.; Papanthasiou, S.; Doerfler, P.A.; Blaine, L.J.; Sun, L.; Yao, Y.; Zhang, C.-Z.; Weiss, M.J.; Pellman, D. Chromothripsis as an on-target consequence of CRISPR-Cas9 genome editing. *Nat. Genet.* **2021**, *53*, 895–905. [[CrossRef](#)] [[PubMed](#)]
89. Offringa, R.; van den Elzen, P.J.M.; Hooykaas, P.J.J. Gene targeting in plants using the *Agrobacterium* vector system. *Transgen. Res.* **1992**, *1*, 114–123. [[CrossRef](#)]
90. Bundock, P.; Hooykaas, P.J.J. Integration of *Agrobacterium tumefaciens* T-DNA in the *Saccharomyces cerevisiae* genome by illegitimate recombination. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 15272–15275. [[CrossRef](#)]
91. van Attikum, H.; Hooykaas, P.J.J. Genetic requirements for the targeted integration of *Agrobacterium* T-DNA in *Saccharomyces cerevisiae*. *Nucl. Acids Res.* **2003**, *31*, 826–832. [[CrossRef](#)]
92. Mladenov, E.; Iliakis, G. Induction and repair of DNA double strand breaks: The increasing spectrum of non-homologous end joining pathways. *Mutat. Res. Fundam. Mol. Mech. Mutagen.* **2011**, *711*, 61–72. [[CrossRef](#)]
93. Decottignies, A. Alternative end-joining mechanisms: A historical perspective. *Front. Genet.* **2013**, *4*, 48. [[CrossRef](#)] [[PubMed](#)]
94. Ceccaldi, R.; Rondinelli, B.; D’Andrea, A.D. Repair pathway choices and consequences at the double-strand break. *Trends Cell Biol.* **2016**, *26*, 52–64. [[CrossRef](#)] [[PubMed](#)]
95. Wang, H.; Xu, X. Microhomology-mediated end joining: New players join the team. *Cell. Biosci.* **2017**, *7*, 6. [[CrossRef](#)] [[PubMed](#)]
96. Friesner, J.; Britt, A.B. Ku80- and DNA ligase IV-deficient plants are sensitive to ionizing radiation and defective in T-DNA integration. *Plant J.* **2003**, *34*, 427–440. [[CrossRef](#)]
97. Gallego, M.E.; Bleuyard, J.-Y.; Daoudal-Cotterell, S.; Jallut, N.; White, C.I. Ku80 plays a role in non-homologous recombination but is not required for T-DNA integration in *Arabidopsis*. *Plant J.* **2003**, *35*, 557–565. [[CrossRef](#)]
98. van Attikum, H.; Bundock, P.; Lee, L.-Y.; Gelvin, S.B.; Hooykaas, P.J.J. The *Arabidopsis AtLIG4* gene is involved in the repair of DNA damage, but not in the integration of *Agrobacterium* T-DNA. *Nucl. Acids Res.* **2003**, *31*, 4247–4255. [[CrossRef](#)]
99. Li, J.; Vaidya, M.; White, C.; Vainstein, A.; Citovsky, V.; Tzfira, T. Involvement of Ku80 in T-DNA integration in plant cells. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 19231–19236. [[CrossRef](#)]
100. Jia, Q.; Bundock, P.; Hooykaas, P.J.J.; de Pater, S. *Agrobacterium tumefaciens* T-DNA integration and gene targeting in *Arabidopsis thaliana* non-homologous end-joining mutants. *J. Bot.* **2012**. [[CrossRef](#)]
101. Nishizawa-Yokoi, A.; Nonaka, S.; Saika, H.; Kwon, Y.-I.; Osakabe, K.; Toki, S. Suppression of Ku70/80 or Lig4 leads to decreased stable transformation and enhanced homologous recombination in rice. *New Phytol.* **2012**, *196*, 1048–1059. [[CrossRef](#)]
102. Vaghchhipawala, Z.E.; Vasudevan, B.; Lee, S.; Morsy, M.R.; Mysore, K.S. *Agrobacterium* may delay plant nonhomologous end-joining DNA repair via XRCC4 to favor T-DNA integration. *Plant Cell* **2012**, *24*, 4110–4123. [[CrossRef](#)]
103. Mestiri, I.; Norre, F.; Gallego, M.E.; White, C.I. Multiple host-cell recombination pathways act in *Agrobacterium*-mediated transformation of plant cells. *Plant J.* **2014**, *77*, 511–520. [[CrossRef](#)]
104. Park, S.-Y.; Vaghchhipawala, Z.; Vasudevan, B.; Lee, L.-Y.; Shen, Y.; Singer, K.; Waterworth, W.M.; Zhang, Z.; West, C.E.; Mysore, K.S.; et al. *Agrobacterium* T-DNA integration into the plant genome can occur without the activity of key non-homologous end-joining proteins. *Plant J.* **2015**, *81*, 934–946. [[CrossRef](#)]
105. Yoshihara, R.; Mitomi, Y.; Okada, M.; Shibata, H.; Tanokami, M.; Nakajima, Y.; Inui, H.; Oono, Y.; Furudate, H.; Tanaka, S. Effects of *Arabidopsis Ku80* deletion on the integration of the left border of T-DNA into plant chromosomal DNA via *Agrobacterium tumefaciens*. *Genes Genet. Syst.* **2020**, *95*, 1–10. [[CrossRef](#)]
106. Nishizawa-Yokoi, A.; Saika, H.; Hara, N.; Lee, L.-Y.; Toki, S.; Gelvin, S.B. *Agrobacterium* T-DNA integration in somatic cells does not require the activity of DNA polymerase theta. *New Phytol.* **2021**, *229*, 2859–2872. [[CrossRef](#)]
107. Francis, K.E.; Spiker, S. Identification of *Arabidopsis thaliana* transformants without selection reveals a high occurrence of silenced T-DNA integrations. *Plant J.* **2005**, *41*, 464–477. [[CrossRef](#)]
108. Mysore, K.S.; Kumar, C.T.R.; Gelvin, S.B. *Arabidopsis* ecotypes and mutants that are recalcitrant to *Agrobacterium* root transformation are susceptible to germ-line transformation. *Plant J.* **2000**, *21*, 9–16. [[CrossRef](#)]
109. Faure, D. Is there a unique integration mechanism of *Agrobacterium* T-DNA into a plant genome? *New Phytol.* **2021**, *229*, 2386–2388. [[CrossRef](#)] [[PubMed](#)]

110. Nisa, M.; Bergis, C.; Pedroza-Garcia, J.-A.; Drouin-Wahbi, J.; Mazubert, C.; Bergounioux, C.; Benhamed, M.; Raynaud, C. The plant DNA polymerase theta is essential for the repair of replication-associated DNA damage. *Plant J.* **2021**, *106*, 1197–1207. [[CrossRef](#)] [[PubMed](#)]
111. Verhage, L. Solving the theta enigma: Polymerase  $\theta$  deficiency causes developmental defects. *Plant J.* **2021**, *106*, 1195–1196. [[CrossRef](#)] [[PubMed](#)]
112. Saika, H.; Shen, Y.; Kim, Y.; Lee, L.-Y.; Gelvin, S.B.; NARO, Tsukuba, Japan and Purdue University, West Lafayette, IN, USA. Unpublished work, 2021.
113. Waterworth, W.M.; Kozak, J.; Provost, C.M.; Bray, C.M.; Angelis, K.J.; West, C.E. DNA ligase I deficient plants display severe growth defects and delayed repair of both DNA single and double strand breaks. *BMC Plant Biol.* **2009**, *9*, 79. [[CrossRef](#)]
114. Bakkeren, G.; Koukolikova-Nicola, Z.; Grimsley, N.; Hohn, B. Recovery of *Agrobacterium tumefaciens* T-DNA molecules from whole plants early after transfer. *Cell* **1989**, *57*, 847–857. [[CrossRef](#)]
115. Singer, K.; Shibolet, Y.M.; Li, J.; Tzfira, T. Formation of complex extrachromosomal T-DNA structures in *Agrobacterium*-infected plants. *Plant Physiol.* **2012**, *160*, 511–522. [[CrossRef](#)] [[PubMed](#)]
116. Singer, K.; Lee, L.-Y.; Gelvin, S.B.; Purdue University, West Lafayette, IN, USA. Unpublished work, 2021.