

Annual Review of Genetics

Integration of *Agrobacterium* T-DNA into the Plant Genome

Stanton B. Gelvin

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

Annu. Rev. Genet. 2017. 51:195–217

First published as a Review in Advance on August
30, 2017

The *Annual Review of Genetics* is online at
genet.annualreviews.org

<https://doi.org/10.1146/annurev-genet-120215-035320>

Copyright © 2017 by Annual Reviews.
All rights reserved

Keywords

DNA repair, DNA recombination, nonhomologous end-joining, NHEJ, plant genetic transformation, chromatin, transgene, agricultural biotechnology

Abstract

Agrobacterium strains transfer a single-strand form of T-DNA (T-strands) and Virulence (Vir) effector proteins to plant cells. Following transfer, T-strands likely form complexes with Vir and plant proteins that traffic through the cytoplasm and enter the nucleus. T-strands may subsequently randomly integrate into plant chromosomes and permanently express encoded transgenes, a process known as stable transformation. The molecular processes by which T-strands integrate into the host genome remain unknown. Although integration resembles DNA repair processes, the requirement of known DNA repair pathways for integration is controversial. The configuration and genomic position of integrated T-DNA molecules likely affect transgene expression, and control of integration is consequently important for basic research and agricultural biotechnology applications. This article reviews our current knowledge of the process of T-DNA integration and proposes ways in which this knowledge may be manipulated for genome editing and synthetic biology purposes.



ANNUAL REVIEWS Further

Click here to view this article's
online features:

- Download figures as PPT slides
- Navigate linked references
- Download citations
- Explore related articles
- Search keywords

Inari Exhibit 1065
Inari v. Corteva

Tumor-/root-inducing plasmid (Ti/Ri plasmid):

a large plasmid conferring virulence upon *Agrobacterium tumefaciens*, *Agrobacterium rubi*, *Agrobacterium vitis*, and *Agrobacterium rhizogenes* strains

Transfer DNA (T-DNA):

a component of tumor-/root-inducing plasmids processed and transferred as single-strand DNA to plants

vir: virulence genes, which—along with their encoded Vir proteins—mediate transformation and are specified by tumor-/root-inducing plasmids; some Vir proteins are transferred to plants

VirD2: Virulence protein involved in processing transfer DNA from the tumor-inducing plasmid, guiding it through the type IV secretion system, and targeting single-strand transfer DNA to plant nuclei

T-strand:

a single-strand form of transfer DNA introduced from *Agrobacterium* into plants

Type IV secretion system (T4SS):

a protein secretion system through which *Agrobacterium* transfers Virulence effector proteins and transfer DNA to plants

INTRODUCTION

Virulent strains of the genus *Agrobacterium* can, in nature, genetically transform a wide variety of plant species. In the laboratory, they can also be used to transform yeast, fungal, and animal cells (64). Although a few other α -proteobacterial species can transfer DNA to plants (10, 49, 62, 139), the high efficiency of *Agrobacterium*-mediated transformation has rendered this process an important tool for plant basic research and a base technology for the agricultural biotechnology industry.

Agrobacterium-mediated transformation is composed of numerous stages (**Figure 1**; 36). All virulent *Agrobacterium* strains harbor tumor-inducing (Ti) or root-inducing (Ri) plasmids, including, for example, *Agrobacterium tumefaciens*, *Agrobacterium vitis*, and *Agrobacterium rubi*; or *Agrobacterium rhizogenes*, respectively. Ti/Ri plasmids encode virulence (*vir*) genes that are induced by plant-released phenolics (112). *vir* gene induction can occur prior to or simultaneously with attachment of the bacteria to plant cells. Following induction, transfer DNA (T-DNA) is processed from the Ti/Ri plasmid through the concerted activity of VirD1 and VirD2 (23). Transfer of a single-strand form of T-DNA (T-strand) and Virulence (Vir) effector proteins is effected through a series of steps: transfer of T-strands (piloted by the covalently attached protein VirD2) and several Vir effector proteins to the plant through an induced Type IV secretion system (T4SS), trafficking of T-strands and Vir protein/plant protein complexes through the plant cytoplasm to the nucleus, T-strand nuclear import, T-DNA integration into the plant genome, and expression of T-DNA-encoded transgenes. Although transient expression of a T-DNA-encoded transgene can occur without integration of T-DNA into host cell chromosomes, stabilization of transgene expression (stable transformation) requires T-DNA integration. This review focuses on these latter steps in stable transformation: T-DNA integration and stable transgene expression. Several recent articles review other aspects of *Agrobacterium*-mediated plant transformation (35–37, 40, 145).

GENERATION OF T-STRANDS AND THEIR ENTRY INTO THE HOST CELL NUCLEUS

Wounding of plant tissue induces the production of numerous defense compounds, including some phenolics that can, along with neutral and acidic sugars, trigger a two-component sensory-response system (VirA–VirG) that induces *Agrobacterium vir* gene expression (140). Among the induced Vir proteins are VirD1, a helicase, and VirD2, a nuclease (45). Together, they nick the T-DNA region of Ti/Ri plasmids, usually between bases 3 and 4 of nearly identical 25-bp left border (LB) and right border (RB) repeat sequences that flank and define T-DNA (143). T-strands are generated as a result of VirD2 nicking at both the LB and the RB. The resulting T-DNA generally contains bases 4–25 of the LB and bases 1–3 of the RB. During the nicking process, VirD2 covalently attaches to the 5' end of T-DNA at the RB (134). VirD2/T-strands exit the bacterium through a T4SS that also exports several Vir effector proteins (VirD5, VirE2, VirE3, and VirF from *A. tumefaciens*; GALLS-CT and GALLS-FL instead of VirE2 from some *A. rhizogenes* strains) (48). Within the plant cell the single-strand DNA-binding protein VirE2 likely noncovalently coats T-strands to form a VirD2/VirE2/T-DNA T-complex (50). Although such a complex has never been isolated, indirect evidence suggests its existence: Fewer T-strands can be isolated from plant cells infected by a *virE2*-mutant *A. tumefaciens* strain than from cells infected by a wild-type *A. tumefaciens* strain (144), and the few T-strands that do integrate into the infected plant genome are greatly truncated, suggesting a lack of protection by VirE2 (97).

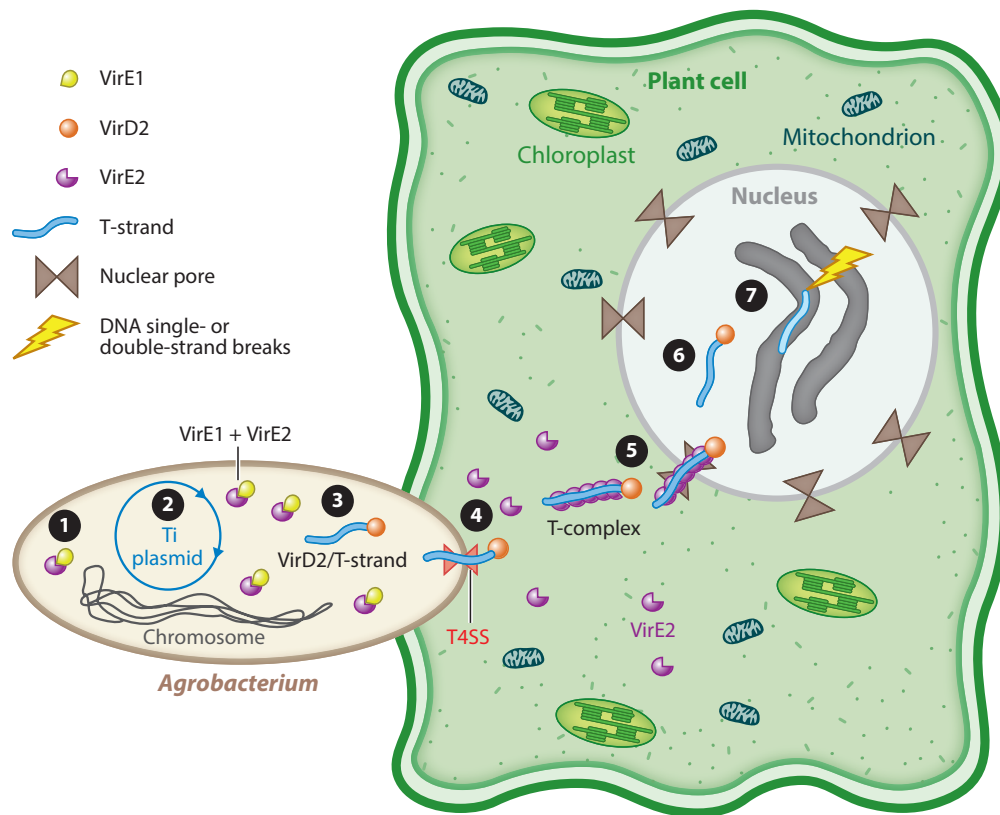


Figure 1

Simplified model of the *Agrobacterium*-mediated plant transformation process emphasizing T-DNA integration and stable transgene expression. Many components of the model remain speculative (e.g., the presence of T-complexes in the plant and where in the plant cell VirE2 may be removed from T-strands). From *left to right*, the hypothesized steps include: (1) Plant phenolic and sugar molecules are perceived by the VirA–VirG two-component regulatory system, resulting in the induction of *vir* genes and the production of Vir proteins. (2) VirD1 and VirD2 nick the T-DNA region of the Ti plasmid at border repeat sequences. (3) VirD2 covalently attaches to the 5' end of one T-strand, forming a VirD2/T-strand complex. (4) VirD2/T-strands and Vir effector proteins exit the bacterium using a T4SS. Although VirE2 associates with its chaperone VirE1 in the bacterium, only VirE2 is transported to the plant. (5) Within the plant, VirE2 protein may associate with T-strands, and the T-complex traffics through the plant cytoplasm and enters the nucleus through nuclear pores. Not shown are other *Agrobacterium* and plant proteins that likely associate with T-complexes. (6) VirE2 protein dissociates from T-strands. (7) T-strands integrate into plant chromosomes, where the encoded transgenes may be expressed. Abbreviations: T-complex, VirD2/T-strand/VirE2 complex; T-DNA, transfer DNA; Ti plasmid, tumor-inducing plasmid; T-strand, a single-strand form of transfer DNA; T4SS, type IV secretion system; *vir*, virulence gene; Vir, Virulence protein.

VirD2 contains a strong C-terminal nuclear localization signal sequence that likely leads T-strands to the nucleus (46). Whether, once within the nucleus, VirD2 targets T-DNA to the plant genome for integration is an open question. Early experiments indicated that VirD2 interacts with TATA-binding protein *in vitro* and in *Arabidopsis* cells (6). Because some early experiments suggested that T-DNA integrates preferentially into promoter regions of transcriptionally active genes (1, 11, 13, 100, 116), some scientists believed that VirD2 was responsible for targeting T-strands for preferential integration into these regions. However, T-DNA does not preferentially target specific regions of the genome for integration (discussed in the next section); thus, it is not likely that proteins involved in transcription are responsible for directing T-DNA to specific chromosomal regions.

Stable transformation: introduction of transfer DNA to plant nuclei followed by integration into the host genome

VirE2: Virulence protein likely involved in protecting single-strand transfer DNA from degradation in plant cells

T-complex: complex formed by single-strand transfer DNA covalently bound by VirD2 protein and noncovalently coated with VirE2 protein

T-DNA junctions: links between integrated T-DNA and plant DNA

T-DNA TARGETING TO THE HOST CHROMOSOMES

Numerous studies have examined regions of the genome into which T-DNA has integrated (1, 3, 13, 51, 54, 55, 100, 102, 115, 119). In these studies, T-DNA/plant DNA junctions from transgenic *Arabidopsis*, *Brachypodium*, or rice plants were identified and characterized. Although the interpretations of these results varied somewhat, in general these authors concluded that T-DNA preferentially integrated into genes, especially into transcriptionally active regions of the genome, promoters, and regions rich in A and T nucleotides (characteristic of promoters). In addition, the stable transformation frequency of the *Arabidopsis* Chromatin Assembly Factor 1 (CAF-1) mutants *fas1* and *fas2* was higher than that of wild-type plants (28). Because the CAF complex is involved in altering chromatin structure and because these mutants additionally displayed higher levels of homologous recombination (HR), the authors suggested that a more open chromatin conformation in the mutants may expose the DNA, permitting both HR-related proteins and T-DNA greater access to plant DNA. Finally, VirD2 protein interacts in plant or yeast cells with proteins important for DNA and histone modification, such as histones (142), S-adenosyl-L-homocysteine hydrolase (involved in DNA methylation) and a MYST-like histone acetyltransferase 2 (66). VirD2/T-DNA may therefore form complexes with enzymes that modify chromatin. Collectively, these results suggested that T-DNA preferentially integrates into euchromatic regions of the genome or regions depleted of nucleosomes.

However, most of the T-DNA junction studies were conducted on regenerated plants that had undergone selection for the expression of transgene resistance markers. Indeed, regeneration of *Agrobacterium*-infected plant tissue without selection indicated that many recovered plants contained silenced transgenes (25, 31). If T-DNA had integrated into a transcriptionally inactive region of the genome (e.g., centromeres, telomeres, heterochromatin), the selection marker may not have been expressed and transgenic plants exposed to selection conditions would not have been recovered. Such selection bias was tested by Kim et al. (58), who recovered T-DNA/plant DNA junctions from infected *Arabidopsis* cells that had not undergone selection for the expression of transgenes. Characterization of these junctions revealed that they reflected the unbiased makeup of the *Arabidopsis* genome. Approximately 10% of that genome comprises highly repetitive DNA, and approximately 10% of the recovered junctions mapped to highly repetitive DNA. Similarly, T-DNA insertions into centromeric, telomeric, and ribosomal DNA (underrepresented in previously characterized selected libraries) were common in the nonselected library, reflecting the proportion of the genome composed of these elements. Microarray and methylation analyses of the preinsertion sites in the nonselected library indicated that they reflected the transcriptional activity of the average gene and that these sites were not hypomethylated. These results indicate that T-DNA integration does not preferentially occur in transcriptionally active or hypomethylated regions of the genome. Thus, previous observations that T-DNA preferentially integrated into active, open chromatin likely resulted from a selection bias placed on the analyzed tissue.

Although T-DNA integration does appear to be random, proteins involved in chromatin modification likely play some role in integration. Mutation or downregulation of *Arabidopsis* and *Nicotiana benthamiana* genes encoding specific histones, histone deacetylases and acetyltransferases, and the histone H3 chaperone SGA1 (ASF1) results in both lower frequency of stable transformation and decreased T-DNA integration into the genome (5, 18, 86, 118, 148). Conversely, overexpression of certain histones can increase transformation susceptibility (53, 86, 118). Thus, proteins important for more general aspects of chromatin architecture, rather than proteins important for transcription of specific genes, may play a role in T-DNA integration (38).

The single-strand DNA-binding protein VirE2 may also play a role in targeting T-strands to chromatin. VirE2 interacts with the bZIP transcription factor VirE2 interacting protein 1 (VIP1)

(123), and VIP1 can interact with histones in vitro (73). Lacroix et al. (63) showed that VIP1 could mediate the interaction of VirE2 with mononucleosomes in vitro. VIP1 could also mediate the association of VirE2–single-strand DNA complexes with mononucleosomes. These results suggest that VIP1/VirE2/T-strand complexes may target nucleosomes comprising chromatin in the nucleus. However, Shi et al. (106) recently showed that VIP1 is not required for *Agrobacterium*-mediated plant transformation, a finding that suggests alternative routes for T-strand targeting of host chromosomes. Indeed, efficient VirE2-independent transformation can occur in the presence of the GALLS protein (48).

ASSAYS FOR T-DNA INTEGRATION

Unfortunately, no in vitro T-DNA integration assay currently exists. Ziemienowicz et al. (149) showed that VirD2 could not ligate single-strand DNA to a model target sequence, indicating that VirD2 does not possess generalized ligation activity. VirD2 could mediate ligation of T-DNA border sequences to themselves, effectively reversing the original nicking reaction (93). Plant extracts could ligate T-DNA to model target sequences in an ATP-dependent manner, suggesting that T-DNA integration depends on plant ligases (149). However, mutation of neither *Arabidopsis* DNA ligase IV nor the plant-specific DNA ligase VI significantly reduced stable *Agrobacterium*-mediated transformation (94, 129). Mutation of DNA ligase VI actually increased transformation (94). These results indicate that in *Arabidopsis*, either DNA ligase I or some yet unknown ligase is responsible for joining T-DNA to plant DNA. It is possible, however, that T-strands that enter the plant cell are not directly ligated to breaks in the plant genome; rather, T-strands may be copied by a host DNA polymerase, and the copy may be ligated to the genome during replication (130; **Figure 2**).

Scientists have often used the frequency of stable *Agrobacterium*-mediated transformation as a surrogate for examining T-DNA integration (e.g., 79). Certainly, the generation of stable transformation phenotypes (e.g., resistance to antibiotics or herbicides, stable green fluorescent protein or β -glucuronidase expression) requires T-DNA integration and stable expression of transgenes. However, as discussed above, T-DNA integration can occur in the absence of transgene expression (25, 31), and integrated transgenes often express at greatly differing levels (e.g., 27, 43, 83, 89). Thus, lack of stable transformation (a transgene expression phenotype) does not necessarily indicate lack of T-DNA integration (a physical genotypic event). Indeed, mutation of the *Arabidopsis* *PARP1* gene results in greatly increased T-DNA integration, but it does not affect stable transformation frequency. This apparent paradox was resolved when Park et al. (94) showed high levels of transgene DNA methylation in *parp1* mutant but not in wild-type plants, suggesting silencing of selection marker genes.

Several physical assays have been developed to investigate T-DNA integration in the absence of selection for transgene expression. In these assays, plant tissue is infected by *Agrobacterium* and, after several days of cocultivation followed by killing the bacteria using antibiotics, the plant tissue is grown in the absence of selection for transgene expression. DNA is subsequently isolated and analyzed for T-DNA integration. These analyses include DNA blot analysis of high-molecular-weight plant DNA that uses a transgene as the hybridization probe (4, 18, 85, 86, 94), PCR using primers directed to T-DNA genes (4, 5, 94), thermal asymmetric interlaced-PCR (TAIL-PCR) to amplify T-DNA/plant DNA junctions (130), and cloning of T-DNA/plant DNA junctions (58). DNA blot-based hybridization assays are especially useful because they can reveal integrated T-DNA fragments that may be missed by conventional or TAIL-PCR. These assays have shown that several genes involved in chromatin formation and histone modification, encoding *Arabidopsis* histone H2A-1, the histone deacetylases HDT1 and HDT2, the histone H3 chaperone SGA1

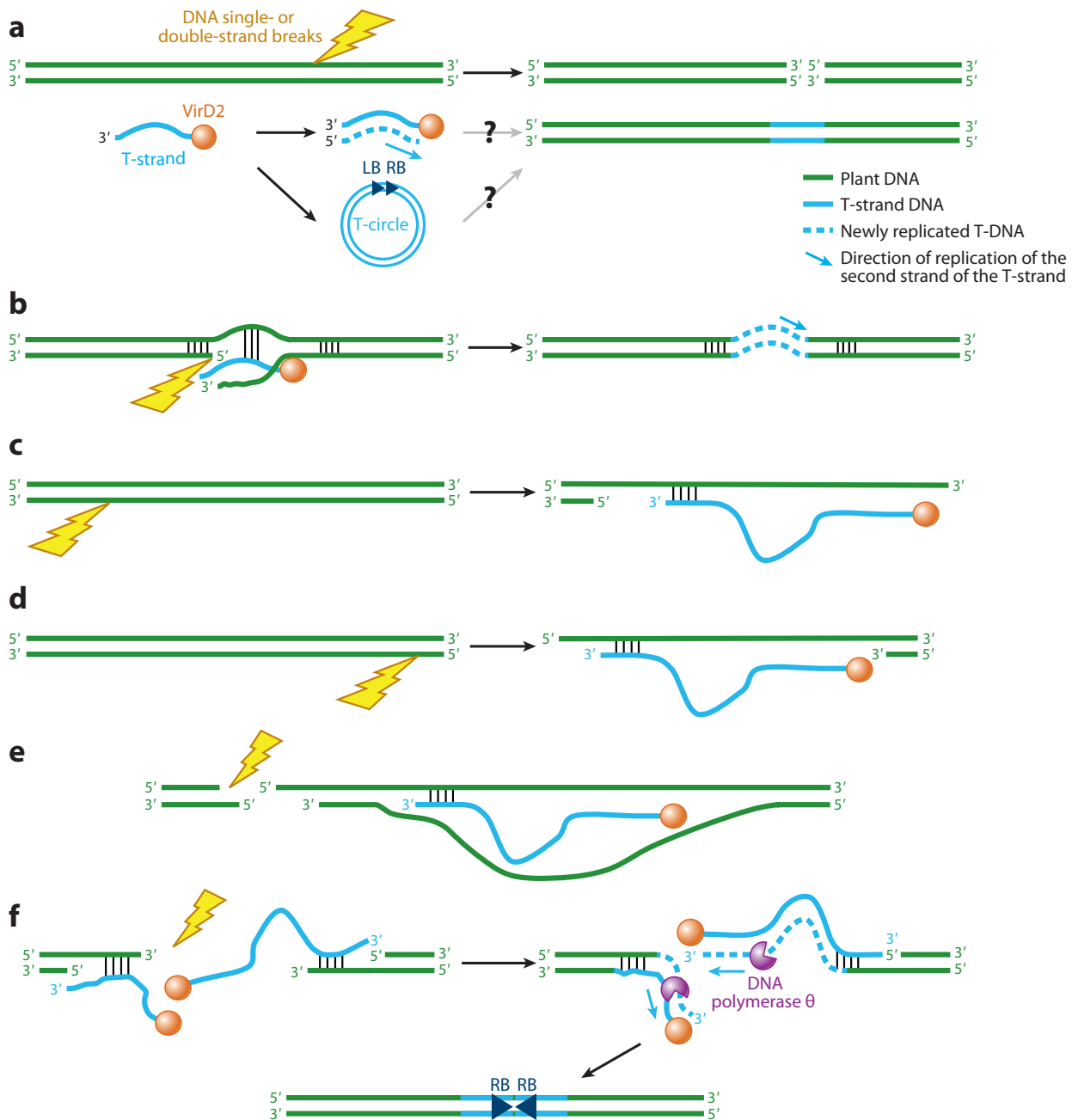


Figure 2

Models for T-DNA integration. (a) Integration of linear or circular double-strand T-DNA molecules into a double-strand break in plant DNA. Currently, there is no evidence (*question marks*) to indicate that T-circles or linear double-strand T-DNA can integrate into plant DNA. (b) Strand invasion of a T-strand into nicked plant DNA. (c,d) Hybridization by microhomology of T-strand sequences with plant DNA after a double-strand DNA break leaves (c) a 3' overhang, or (d) a 5' overhang. (e) Strand invasion of a T-strand into plant DNA following a double-strand DNA break. (f) Copying of T-strand sequences by DNA polymerase θ (purple) into a plant double-strand DNA break. Shown are two T-strand copies in a head-to-head conformation. Abbreviations: LB, left border; RB, right border; T-circles, extrachromosomal circular transfer DNA molecules; T-DNA, transfer DNA; T-strand, a single-strand form of transfer DNA; Vir, Virulence protein. Models c–e are adapted from an unpublished slide of Dr. Seiichi Toki and used with permission. The model shown in panel f is an extension of that in van Kregten et al. (130).

(ASF1), and tobacco histone H3, are important for T-DNA integration. DNA blot assays have also been used to show that mutation of DNA repair and recombination genes encoding XRCC4, Ku70, Ku80, DNA ligase VI, and PARP1 results in increased T-DNA integration into the plant genome (94).

PATTERNS OF T-DNA INTEGRATION AND EFFECTS ON TRANSGENE SILENCING

T-DNA integration into the plant genome may occur at one or several loci, which may be segregated in subsequent generations. The copy number of integrated T-DNAs varies among studies and within individual studies from a single copy to tens of copies. Clearly such copy number differences do not always result from the use of different *Agrobacterium* strains, plant species, or transformation protocols, because copy number varies even between independent events within a single study. Oltmanns et al. (91) showed that transformation frequency and integrated T-DNA copy number in *Arabidopsis* (generated by flower-dip transformation) and in maize [generated by immature zygotic embryo transformation (30)] did not correlate with the T-DNA binary vector copy number in *Agrobacterium*. However, when T-DNA was integrated into and launched from the *Agrobacterium* chromosome, the transformation frequency was somewhat reduced (twofold in *Arabidopsis*, eightfold in maize) but the integrated T-DNA copy number was greatly reduced to very few multicopy insertions in hundreds of individual events investigated. In addition, these transgenic lines almost always lacked integrated T-DNA binary vector backbone sequences. These authors suggested that launching T-DNA from the *Agrobacterium* chromosome may be useful to reduce the integrated T-DNA copy number in the resulting transgenic events. More recently, Zhi et al. (147) showed that, although the initial maize transformation frequency could vary depending on the type of binary vector used, the percentage of single-copy integration events was not altered using binary vectors with different origins of replication and, hence, bacterial copy number. Additionally, Sivamani et al. (110) used an RNAi approach targeting selectable marker transcripts to obtain transgenic maize plants with single-copy T-DNA insertions.

The importance of single T-DNA integration events lies not only in the relative ease of analyzing these plants but also in the reduced likelihood that single-copy integrants will experience transgene silencing. One of the remaining limitations to exploiting *Agrobacterium*-mediated transformation is variability and stability of transgene expression. Independent transgenic events generated by the same *Agrobacterium* strain frequently show a broad range of transgene expression, from very high to totally silent (43, 89). Plants that express transgenes well under laboratory conditions may undergo reduction in transgene expression under certain field conditions (8, 44). Variability in transgene expression (7, 47, 52, 57, 84, 95) can make it difficult to evaluate transgenic events and usually necessitates the generation and testing of hundreds or, in industrial settings, thousands of transgenic plants to identify those few lines with stable and high-level transgene expression.

The cause of transgene expression variability remains unclear. However, certain T-DNA integration characteristics often relate to the extent of transgene expression. Multiple T-DNA copies may link at each locus. RB-to-LB linkages (head-to-tail) indicate tandem integrated T-DNAs in a direct repeat orientation, but inverted repeat LB-to-LB (tail-to-tail) or RB-to-RB (head-to-head) integration events may also occur. Head-to-head inverted repeats are common and are often associated with transgene silencing (9, 17).

Although transgene silencing is often associated with multiple T-DNA insertions at the same locus, this need not always be so (65, 104), and single-T-DNA insertions may also silence (27, 80, 83). Indeed, independent transgenic events each harboring single-T-DNA insertions may show a wide range of transgene expression. Because the only known difference among these events is the

T-DNA binary vector: a plasmid containing transfer DNA but not virulence genes

Transgene silencing: lack of expression of a transferred exogenous gene

Microhomology: the presence of identical short DNA sequences found in multiple genetic elements

location of T-DNA insertion, the cause of variation in transgene expression has been attributed to position effects, possibly reflecting the transcriptional permissiveness of the chromatin into which T-DNA has integrated (19, 38, 137).

The linkage of multiple independent T-DNAs, rather than duplication of a single T-DNA at a locus, was shown by infecting plants with two *Agrobacterium* strains containing T-DNAs with different selection markers. Simultaneous or sequential selection of plant tissue on medium containing these selection agents resulted in tissue resistant to both antibiotics and containing both T-DNAs. In some instances, the two different T-DNAs were closely linked in the plant genome (20–22).

T-DNA/PLANT DNA JUNCTIONS

T-DNA has been used to generate large insertional mutation (knock-out) libraries in *Arabidopsis* (1, 102, 115), rice (3, 13, 51, 54, 55, 100), and *Brachypodium distachyon* (119). As a result of sequencing the T-DNA/plant DNA junctions (flanking sequence tags) in these libraries, the sequences of hundreds of thousands of these junctions are known. A recent analysis of T-DNA/plant DNA junctions in the *Arabidopsis* GABI-Kat collection (59) illustrates what has been found in other large T-DNA insertion collections and in other studies of fewer junctions (39, 78). Kleinboelting et al. (59) annotated 11,103 LB and 699 RB junctions. They found that within the T-DNA portion of the junctions, RBs were generally more precise than were LBs, i.e., >45% of RBs were either intact (containing bases 1–3 of the border repeat) or had deletions of four bases or fewer. The median size of the RB deletion was two nucleotides. However, only approximately 25% of LBs were relatively intact (precise or <4 base deletions), and deletions extending from the LB junction were considerably larger, on average, than were deletions extending from the RB junction: The median size of LB deletions was seven nucleotides. The relative precision of the RB junction sequences has been ascribed to VirD2 protein that is covalently linked to the RB, presumably protecting it from exonuclease digestion. Larger LB junction deletions have been ascribed to lack of protection of the LB end region in the VirD2/VirE2/T-strand complex (97). *virE2* mutant *Agrobacterium* strains are extremely attenuated in virulence (113), although they accumulate near-wild-type levels of T-strands within the bacterium (114, 131). This result and the observation that transgenic plants expressing *virE2* can complement a *virE2*[−] *Agrobacterium* strain (15) indicate that the virulence function for VirE2 occurs in the plant. Indeed, fewer T-strands were observed in plant cells soon after inoculation with a *virE2*[−] *Agrobacterium* strain than were seen after infection by a wild-type strain (144). The LB region of T-DNA/plant DNA junctions of the few transformants generated by a *virE2*[−] *Agrobacterium* strain were deleted to a greater extent than were such LB regions of junctions generated by *virE2*⁺ strains (97). These results indicate that VirD2 likely protects T-DNA RB sequences in the plant, whereas lack of VirE2 at the extreme LB region results in truncation of integrated T-strands, either by exo- or endonuclease activity or by revealing T-strand 3' end sequences for microhomology-mediated strand annealing (see The Process of T-DNA Integration below).

T-DNA/plant DNA junctions are generally not clean, i.e., they are not comprised of precisely cleaved T-DNA borders precisely integrating into a plant sequence. Often deletions, insertions, and rearrangements occur at the junctions. One of the earliest sequence analyses of T-DNA/plant DNA junctions identified internal T-DNA sequences at the LB junction (108). T-DNA binary vector sequences (i.e., the plasmid backbone for T-DNA vectors) frequently are transferred and integrated into the plant along with T-DNA (61, 77, 138). The integration of vector backbone sequences may result from border skipping by the VirD2 endonuclease (61) or initiation of DNA transfer from VirD2–LB sequences (26). Because sequences adjacent to the T-DNA region were

missing from most transgenic *Arabidopsis* and maize plants generated by launching T-DNA from the *Agrobacterium* chromosome, Oltmanns et al. (91) speculated that the transfer of plasmid backbone sequences was a consequence of the relatively high copy number of the T-DNA binary vectors used for most experiments. In addition to plasmid backbone sequences, *Agrobacterium* chromosomal DNA fragments have been detected at T-DNA/plant DNA junctions in approximately 0.4% of the examined lines of large T-DNA insertion collections (126). The mechanism of *Agrobacterium* chromosomal DNA transfer to plants remains unknown.

Filler DNA sequences frequently occur between T-region borders and the linked plant DNA and between multiple linked T-DNA molecules (either direct or inverted T-DNA repeats). Although the origin of these sequences was not always apparent in early studies (e.g., 141), recent improvements in DNA sequencing and informatics analyses have yielded more information. In addition to internal T-DNA sequences and *Agrobacterium* chromosomal sequences previously mentioned, most filler sequences map to plant DNA. Of the 1,508 filler sequences analyzed by Kleinboelting et al. (59), almost 95% mapped to the *Arabidopsis* chromosome. A large percentage (approximately 36%) mapped to the same chromosome as the T-DNA insertion, and 3.1% mapped to within 100 nucleotides of the T-DNA insertion site. Other filler sequences randomly mapped to the other four *Arabidopsis* chromosomes. The source of the filler sequences frequently showed microhomology with the plant target sequence at the insertion site. This finding, along with the observation that the T-DNA sequence at the insertion site frequently shows microhomology with the plant preintegration target site, suggests a microhomology-mediated mechanism of T-DNA insertion (see The Process of T-DNA Integration below).

Other plant chromosomal abnormalities are also frequently found at T-DNA/plant DNA junctions. When both junctions of an integrated T-DNA are sequenced, deletions of plant DNA are commonly seen (59). Deletions are also often seen in T-DNA insertion lines in regions of the genome other than T-DNA integration sites (e.g., 136). However, it is not clear whether these deletions result from T-DNA integration (or perhaps abortive integration, whereby the integration process is interrupted before T-DNA is fully integrated), or from plant tissue culture processes (136). In addition, chromosomal inversions and translocations are commonly associated with T-DNA insertions (59, 76, 98, 117). Indeed, Clark & Krysan (16) showed that 19% of the examined T-DNA insertion lines from the Salk collection have chromosomal translocations. These findings suggest that multiple double-strand DNA breaks may be associated with T-DNA integration.

PROTEINS IMPORTANT TO T-DNA INTEGRATION

The screening of T-DNA insertion libraries (148) and virus-induced gene silencing (VIGS) libraries (5) identified numerous genes important for *Agrobacterium*-mediated transformation of *Arabidopsis* and *N. benthamiana*. The *Arabidopsis* mutants were designated *rat* (resistant to *Agrobacterium*-mediated transformation), and the *N. benthamiana*-silenced lines were designated ACG (altered crown gall). Some of these lines retained high susceptibility to transient transformation (87). The inability of these lines to stabilize transgene expression suggested either that they were deficient in T-DNA integration, or that T-DNA had integrated but the encoded transgenes were silenced. Biochemical assays for T-DNA integration into high-molecular-weight plant DNA, as described above, indicated that several genes were important for T-DNA integration, including those encoding histone H2A (86), histone H3 (5, 148), a nodulin-like protein (NLP) (5), the histone H3 chaperone SGA1 (18), the histone deacetylases HDT1 and HDT2 (18), and the NOT domain transcription factor VIP2 (4). Many of the genes encode chromatin components or chromatin-modifying enzymes, and it is interesting that *vip2* mutants have altered levels of histone and histone deacetylase transcripts (4).

Filler DNA: DNA sequences at transfer DNA/plant DNA junctions that do not derive from those at the insertion site

Transient transformation: introduction of transfer DNA to plant nuclei without further integration

In many ways, T-DNA integration resembles DNA break repair through nonhomologous end-joining (NHEJ) or an alternative microhomology-mediated end-joining (MMEJ) process. In all instances, double-strand DNA breaks may be misrepaired, resulting in deletions, insertions, rearrangements, and even translocations at the break sites (24, 82). Salomon & Puchta (101) showed that cleavage of an I-*SceI* site artificially introduced into the tobacco genome could result in many of these types of mutations at the site, including the appearance of filler DNA sequences containing limited (micro)homology with the sequences surrounding the cut site. Indeed, these authors found that T-DNA (used to introduce the I-*SceI* enzyme) occasionally integrated at this site. These findings were confirmed by Chilton & Que (14) and Tzfira et al. (121), who used, respectively, I-*CeuI* and I-*SceI* to cut the corresponding sites introduced into tobacco. Overall, creating these specific double-strand DNA breaks resulted in integrating T-DNA at these sites in 1–3% of the examined transgenic events, a very high number relative to possible random integration into the tobacco genome. X-ray irradiated plant protoplasts that have increased double-strand DNA breaks also show higher transformation rates, suggesting that DNA double-strand breaks (or induction of damage-repair enzymes) enhance foreign DNA integration (60).

Because of the perceived similarities between T-DNA integration and DNA break repair and recombination processes, several laboratories have tested *Arabidopsis* or rice DNA repair/recombination mutants for alterations in stable *Agrobacterium*-mediated transformation. T-DNA need not have homology to target integration sites; indeed, gene targeting is highly inefficient in plants (42). Therefore, NHEJ has been the favored mechanism by which T-DNA integrates. During NHEJ, double-strand DNA breaks are repaired by a major classical (cNHEJ) pathway involving recognition and alignment of the broken ends by the Ku70/Ku80 heterodimer, polishing of ragged ends, and ligation by the DNA ligase IV/XLF/XRCC4 complex. Along with cNHEJ, or when the cNHEJ pathway is disrupted, an alternative (aNHEJ) or backup pathway is activated that uses histone H1 and PARP1 to recognize and align broken ends. XRCC1 and DNA ligase III may be used in animals to ligate the ends (24, 82), and DNA ligase I or the plant-specific DNA ligase VI (135) may be used in plants.

Whether T-DNA integration uses either of these pathways remains in dispute. Indeed, Park et al. (94) showed that simultaneously disrupting *ku80* (involved in cNHEJ) and *parp1* (involved in aNHEJ) had no negative effect on either stable transformation or T-DNA integration into the *Arabidopsis* genome. Additionally, Mestiri et al. (79) showed that simultaneous disruption of up to four DNA repair and recombination pathways, including cNHEJ and aNHEJ, greatly decreased but did not eliminate transformation. Thus, although some proteins associated with DNA repair and recombination may influence T-DNA integration, none of the pathways are essential.

Several laboratories have investigated the effects of mutating individual genes in the cNHEJ pathway on stable *Agrobacterium*-mediated transformation, and they have come to opposing conclusions regarding their importance. For example, van Attikum et al. (128) showed that several NHEJ genes are important for T-DNA integration in yeast, but the gene encoding DNA ligase IV is not important for stable transformation of *Arabidopsis* (129). Friesner & Britt (32) showed that *Arabidopsis ku80* and *lig4* mutants showed slightly reduced stable transformation frequencies; a similar decrease in stable transformation was observed by Li et al. (68). However, Gallego et al. (34) did not see an effect of a *ku80* mutation on stable *Arabidopsis* transformation. Jia et al. (56) did not see any major effect of *Arabidopsis ku* mutations on root or flower dip transformation. Mestiri et al. (79) observed a 1.5- to 2.5-fold decrease in *Arabidopsis* transformation using *ku80*, *xrcc1*, *xpf*, or *xrcc2* mutant plants and a greater decrease with double, triple, or quadruple mutants. Finally, rice *ku* and *lig4* mutants showed decreased *Agrobacterium*-mediated transformation (90, 99). The differences seen by these laboratories may reflect the use of different species (rice versus *Arabidopsis*) and transformation methods (somatic-cell versus flower dip-gametophytic). However, none

of these laboratories directly investigated the importance of these NHEJ genes on T-DNA integration; rather, they inferred T-DNA integration from stable transformation, taking into account active insertions only.

Recently, two laboratories investigated the importance of *XRCC4* and other NHEJ genes on both the active insertions of stable transformation and the active and silenced insertions of T-DNA integration (94, 127). Surprisingly, mutation of many cNHEJ genes increased stable transformation and T-DNA integration. In no instance did mutation of any individual NHEJ or other DNA repair/recombination gene decrease stable transformation or T-DNA integration. Similarly, simultaneous mutation of up to four genes in different DNA repair/recombination pathways increased both stable transformation and T-DNA integration (Y. Shen, Y. Kim, L.-Y. Lee, A. Kaundal, K.S. Mysore & S.B. Gelvin, unpublished data). Mutation of *PARP1* (involved in aNHEJ) increased T-DNA integration without increasing stable transformation, suggesting that mutation of *PARP1* silences integrated transgenes. Indeed, an integrated *gusA* transgene is highly methylated in a *parp1* mutant (94).

Recently, van Kregten et al. (130) showed that mutation of the *Arabidopsis* gene encoding DNA polymerase θ [TEBICHI (*teb*) mutants] almost completely inhibited stable, but not transient, transformation of *Arabidopsis*. It is likely that DNA polymerase θ mediates the first step of T-DNA integration.

TEBICHI (*TEB*):
an *Arabidopsis* gene
encoding DNA
polymerase θ

THE PROCESS OF T-DNA INTEGRATION

Several mechanisms have been proposed for T-DNA integration. In general, these can be defined as T-strand invasion (aided by microhomology of T-strands with the plant double-strand chromosome target site) or conversion of T-strands to double-strand molecules followed by ligation into plant double-strand DNA breaks (99, 122). These pathways are not necessarily mutually exclusive, and variants of these mechanisms can be envisaged (67, 99, 122, 130; **Figure 2**).

T-DNA enters plants cells as a single-strand molecule (120, 144) capped on its 5' terminus by VirD2 protein. VirE2 protein may coat T-strands, but this interaction may not protect the extreme 3' T-strand terminus. In addition, VirE2 would have to be stripped from T-strands prior to or during T-DNA integration (74, 124, 146), unmasking T-strand sequences and rendering them sensitive to nuclease degradation. Removal of VirE2 protein would also provide substrates for homology-based (or microhomology-based) matching of T-strand sequences with complementary plant target DNA sequences. Complementary pairing of the T-strand would likely not occur exactly at the T-strand 3' end, and T-strand sequences 3' of the homology site would be trimmed during further steps in integration. Either nuclease degradation or trimming of T-strand 3' ends following homology-based strand invasion would result in integrated T-DNA molecules lacking sequences near the 3' end (LB region), and this is what is frequently seen at T-DNA/plant DNA junctions. Thus, the homology-based strand invasion model of integration garnered initial favor (39, 78).

However, the strand invasion model cannot readily account for the frequently found head-to-head (RB-RB) or tail-to-tail (LB-LB) integrated T-DNA dimers. T-strands would first have to be converted to a double-strand form before ligation. Thus, the formation of double-strand T-DNA molecules in the plant nucleus prior to ligation into plant double-strand DNA breaks has become the favored mechanism of T-DNA integration. Evidence for double-strand T-DNAs in the nucleus is abundant: Narasimhulu et al. (88) designed a T-DNA that would create a T-strand containing the noncoding strand of a *gusA*-intron gene. This T-strand would have to be converted into a double-strand molecule in order to express the *gusA* gene. β -Glucuronidase (GUS) activity could be detected in plant cells, even when T-DNA integration was greatly diminished because

T-circles:

extrachromosomal
circular transfer DNA
molecules found in
plant cells after
Agrobacterium
infection

T-DNA borders:

similar 25-bp directly
repeated sequences
that flank and define
the transfer DNA
region of tumor- or
root-inducing
plasmids on the right
and left

of a mutation in VirD2 protein (85) or because of mutation of a histone gene important for integration (86). In addition, several laboratories have used site-specific recombinases to integrate T-DNA into cognate sites placed in plant genomes (125, 132, 133). These recombinases require double-strand DNA as a substrate.

The potential importance of double-strand DNA breaks as targets for T-DNA integration was implied by experiments in which rare-cutting enzymes were used to target T-DNA integration into these sites (14, 101, 121), as discussed above. However, to ligate double-strand T-DNA molecules into these (or any) sites, T-strands would first need to be converted to double-strand molecules. Liang & Tzfira (71) presented evidence that T-strands, delivered by *Agrobacterium*, could be converted to a double-strand form using synthetic DNA or RNA oligonucleotides introduced into plant cells. They speculated that such oligonucleotides naturally found in plant nuclei could also serve this purpose. These authors additionally suggested that deletion of T-DNA sequences near LBs could occur as a result of primers hybridizing to T-strands at internal positions. Double-strand DNA synthesized as a result of replicative extension from such primers would lack the 3' end of the T-strand, and the single-strand ends would be subject to exonuclease removal.

Conversion of T-strands to double-strand circular molecules was demonstrated by Singer et al. (109). These authors designed T-DNA binary vectors that incorporated plant- and bacteria-active selection markers, as well as a bacterial origin of replication, within the T-DNA region. *Agrobacterium* harboring these binary vectors was used to infect *N. benthamiana* plants, and total DNA was isolated from the infected tissue. The DNA was transformed into *E. coli*, and plasmids (either the original binary vector or T-circles generated in planta) were rescued. Screening against the binary vector backbone antibiotic resistance gene excluded binary vector molecules rescued from contaminating *Agrobacterium* cells, and the regions spanning T-circle T-DNA borders were sequenced. These T-circle border junctions had all the properties of plant DNA junctions of integrated T-DNAs. Many had complex structures of T-DNA multimers, including tandem and inverted T-DNA repeats, and many used microhomology at the joined borders that sometimes resulted in deletions, additions, and other rearrangements. The authors also found filler DNA and T-DNA binary vector sequences at T-circle border junctions. It is not clear what role these T-circles play in the process of *Agrobacterium*-mediated transformation. The authors speculated that they could be dead-end intermediates that do not integrate or intermediates that serve as a substrate for integration. They further inferred that by investigating the formation of these T-circles, scientists could gain knowledge of the mechanism of T-DNA integration.

The finding that DNA polymerase θ is essential for T-DNA integration suggests a mechanism for integration (130). The authors inferred that DNA polymerase θ captures the 3' end of T-strands at plant double-strand breaks through microhomology between sequences at or near these 3' ends and plant sequences near the break sites. Capture of two T-DNAs at each broken plant DNA end could generate a substrate for head-to-head T-DNA dimers, although there is no easy explanation for the generation of tail-to-tail or head-to-tail dimers. The propensity for DNA polymerase θ to primer-template switch can also account for filler DNA sequences at T-DNA/T-DNA and T-DNA/plant DNA junctions. This model of DNA polymerase θ -mediated T-DNA integration incorporates aspects of both double-strand DNA break generation (in the plant DNA) and microhomology-mediated strand annealing (between T-strands and resected plant DNA molecules) (67). This model can also explain microhomologies frequently found at T-DNA LB and plant target DNA sequences (e.g., 59), the lack of a requirement for NHEJ proteins for transformation (79, 94), and observations that mutants lacking NHEJ components may undergo slower break repair, providing more opportunity for homology search and subsequent T-strand replication at the break sites (94, 127).

CAN SCIENTISTS CONTROL THE SITE OF T-DNA INTEGRATION?

T-DNA integration is random in the plant genome. Random integration has useful consequences: Scientists have used T-DNA as an insertional mutagen to generate large libraries of *Arabidopsis*, rice, and *Brachypodium* plants with mutations in the majority of known genes (1, 3, 13, 51, 54, 55, 100, 102, 115, 119). However, random T-DNA integration may have negative consequences for transgenic plants used for research and biotechnological purposes. These include (a) disruption of genes or regulatory elements important for plant growth, development, or productivity; (b) integration of transgenes into regions of the genome that are poorly expressed or may show altered or unstable expression when plants are placed in the field or are under stress; and (c) costly and time-consuming regulatory hurdles to characterize T-DNA insertion sites for field release and commercial growth of transgenic plants. Because of these potential problems associated with random T-DNA integration, targeting T-DNA integration into specific well-characterized regions of any plant genome offers numerous advantages for both research and applications.

Targeting transgenes to specific genomic sites is routine in bacteria and fungi, and it is frequently accomplished in mammalian systems. However, efficient knock-in technologies have until recently been a difficult-to-achieve holy grail of plant biology. In general, sequences introduced into plant cells using *Agrobacterium*-mediated transformation or other techniques such as particle bombardment (gene gun) randomly integrate, even if numerous kilobases of homology exist between the introduced DNA and the plant genome. Even with *Arabidopsis*, a plant containing a relatively small genome (approximately 140 Mbp), gene targeting occurred at a frequency of 10^{-4} – 10^{-3} (42).

In the past decade, numerous advances have been made in genome engineering through the design of custom nucleases to create double-strand DNA breaks at specific chromosomal sites (12, 33, 75, 92, 96). The double-strand breaks made by these enzymes either can be misrepaired by NHEJ to generate mutations at specific sites or can stimulate homology-directed repair that introduces new sequences into specific genomic target sites (103). However, the design of these engineered nucleases is cumbersome and generally out of reach for the average laboratory. More recently, CRISPR-Cas9 technology has been used to generate double-strand breaks at specific sites. This system uses the Cas9 nuclease, which is directed to specific DNA target sites by a single-guide RNA (sgRNA); thus, one must only build appropriate sgRNA expression cassettes to generate specific double-strand DNA breaks. Use of this technology is considerably simpler than the generation of custom endonucleases, and it is thus rapidly gaining favor for genome engineering (e.g., 29, 69, 70, 72, 105). Using CRISPR/Cas9 technologies, scientists have been able to insert single genes or gene fragments into specific sites at frequencies greater than 10%.

As promising as CRISPR/Cas9 technology may be, no laboratory has yet integrated a large DNA fragment containing numerous genes into a specific chromosomal site. Insertion of large batteries of genes will be important for synthetic biology purposes, e.g., to introduce an entire biosynthetic pathway into an organism. *Agrobacterium* can transfer large fragments of DNA containing numerous genes [wild-type oncogenic T-DNAs can be >25 kbp; DNA transfer >100 kbp has been reported under laboratory conditions (41, 81)]. Thus, unlike other technologies such as particle bombardment that often fragment long DNA molecules during transformation, *Agrobacterium* can deliver intact DNA molecules large enough to encode numerous genes. Because *Agrobacterium*-mediated DNA transfer represents a natural system, *Agrobacterium* has the propensity to generate a single copy or a low copy number of integration events with defined ends. Thus, it is less disruptive to the plant genome structure. Finally, no special equipment or expensive supplies (such as gold particles) are needed for transformation. Therefore, it is the preferred plant transformation method for laboratories with limited resources. The challenge will be to marry

Homology-directed repair: precise repair of a DNA molecule using a nucleic acid fragment of identical or similar sequence

the DNA transfer ability of *Agrobacterium* with a technology that will target T-DNA to specific chromosomal sites.

As mentioned above, rare-cutting endonucleases (I-*SceI* or I-*CeuI*) have been used to generate site-specific breaks in plant genomes that trap T-DNA molecules for integration (14, 101, 121). The target sites for these nucleases were randomly placed into the plant genome by *Agrobacterium*-mediated transformation. Because double-strand DNA breaks appear to trap T-DNA, a useful approach may be to use CRISPR/Cas9 to direct production of these breaks to specific plant sequences. These sequences should be carefully chosen such that:

1. They reside near or between genes that show predictable and stable expression, even under field stress conditions.
2. They reside near genes not important for plant growth, development, or productivity. Because insertions in single-copy genes may be detrimental to the plant, single-copy genes should not be a target for T-DNA insertion.
3. The sequences into which T-DNA will insert should be unique in the genome to mitigate off-targeting by the CRISPR/Cas9 nuclease system. Such a T-DNA targeting system would be useful for both basic research and for agricultural biotechnology.

CAN SCIENTISTS USE *AGROBACTERIUM* TO DELIVER BUT NOT INTEGRATE T-DNA?

As genome engineering of plants becomes an increasingly important tool, it would be useful, especially for regulatory purposes, to develop technology to prevent integration of DNA encoding gene editing reagents, such as T-DNA encoding sgRNAs and Cas9 protein, into the plant genome (see the sidebar titled Plant Genome Engineering). T-DNAs encoding genome editing tools frequently integrate into the genomes of edited plants and are usually segregated out in subsequent generations. However, this can be time-consuming and may be cumbersome and expensive for vegetatively propagated plant species, including potato, sweet potato, banana, cassava, and sugar cane, in which outcrossing integrated transgenes is difficult. It would therefore be useful to design an *Agrobacterium*-based gene transfer technology whereby T-DNA-encoded genome editing reagents are transiently expressed but not maintained in the resulting plant.

Two possible approaches may be explored to effect transient but not stable *Agrobacterium*-mediated transformation. The first involves DNA polymerase θ . This DNA polymerase is a major player in the first step in T-DNA integration (130). *Arabidopsis* plants mutant in DNA polymerase θ are highly susceptible to transient transformation but resistant to stable transformation. Thus, they can likely be used to shift the DNA processing equilibrium from random T-DNA integration to homology-dependent repair. However, there would have to be some way to decrease transiently

PLANT GENOME ENGINEERING

Plant genome engineering is a rapidly developing field of plant synthetic biology and agricultural biotechnology. *Agrobacterium*-mediated transformation is currently the most commonly used technology to deliver genome engineering reagents, although viral-mediated delivery methods are also being developed. Research needs to focus on development of novel *Agrobacterium*-based methods that will deliver large amounts of DNA substrates, which are needed for efficient homology-directed DNA repair, without ectopic integration of these substrates or of Cas9/sgRNA-encoding genes.

DNA polymerase θ activity of multiple crop species and varieties. This potentially could be accomplished using RNAi approaches to target the *TEB* (DNA polymerase θ) gene orthologs in these species.

A second approach to eliminating T-DNA integration would be to develop an *Agrobacterium* strain that could efficiently deliver T-DNA to the plant nucleus but not permit T-DNA to integrate. Numerous *Arabidopsis* mutations exist that allow such a scenario to occur (148), but this would likely not be a useful approach for crop species. A mutant VirD2 protein has been characterized that mediates somewhat reduced transient transformation (20–25% that of the wild-type protein) but greatly reduced stable transformation and T-DNA integration (1–3% that of wild-type VirD2) (85, 88). This mutant protein contains a substitution of two serines for four amino acids (GRGG) in the penultimate VirD2 ω domain (107). Further manipulation of this mutant protein, or of the ω domain, may result in a VirD2 protein that can transfer T-DNA to plant nuclei with wild-type efficiency but fail to integrate it into the plant genome.

SUMMARY POINTS

1. *Agrobacterium*-mediated plant genetic transformation is a complex process involving numerous bacteria- and plant-encoded proteins.
2. T-DNA integration and transgene expression are the final steps in stable transformation.
3. T-DNA integration is random in the plant genome. The integration locus likely affects transgene expression (position effect). The integrated T-DNA conformation (repeated units of T-DNA) may also affect transgene expression.
4. T-DNA integration is rarely a clean event. Deletions, insertions, filler DNA sequences, binary vector backbone sequences, and chromosomal translocations frequently accompany T-DNA integration.
5. Although T-DNA integration resembles DNA repair processes such as NHEJ and MMEJ, integration does not require the same proteins involved in these repair processes.
6. DNA polymerase θ is required for T-DNA integration in *Arabidopsis*.
7. Control of the T-DNA integration site may be important for future plant synthetic biology experiments.
8. Efficient T-DNA delivery without T-DNA integration may be important for future plant genome engineering experiments.

FUTURE ISSUES

We still have much to learn about the mechanism of T-DNA integration into the plant genome. The recent finding that DNA polymerase θ is involved in the first step of integration, at least in *Arabidopsis*, is a major breakthrough but invites numerous other questions:

1. What is the role of DNA damage in selecting sites for integration? Does T-DNA target single-strand DNA nicks or double-strand DNA breaks?

2. If T-DNA targets damaged DNA sites, does *Agrobacterium* take advantage of pre-existing break sites, or does plant infection by *Agrobacterium* induce DNA damage? Recently, Song & Bent (111) observed increased double-strand breaks (measured using comet assays) and phosphorylated histone H2A.X (γ -H2A.X, a sensitive marker for DNA breaks) in *Arabidopsis* infected by *Pseudomonas syringae* pv. *tomato* DC3000 and in tomato and potato infected by various pathogenic fungi. Thus, plant pathogens can cause DNA damage to their hosts. If *Agrobacterium* infection induces DNA breaks, is the plant responding to secreted pathogen-associated molecular patterns or molecules on the bacterial surface, or do transferred Vir effector proteins induce DNA damage? Thus, is the *Agrobacterium*-mediated transformation process inherently mutagenic?
3. If T-strands target DNA break sites, how is that targeting accomplished? Do Vir effector proteins associated with T-strands (VirD2 and VirE2) target break sites directly, or do they interact with plant DNA repair proteins (including γ -H2A.X) associated with these sites?
4. To what extent is VirD2 involved in the integration process? Is VirD2 removed from T-strands prior to integration, or are T-strands copied into plant DNA by DNA polymerase θ ?
5. If T-DNA (in single-strand or double-strand form) is ligated into plant DNA, which DNA ligase or ligases are responsible for this activity?
6. Can we specify the T-DNA integration site in the plant genome, and can we additionally control the integration process to generate high-quality integration events indicated by a high, stable transformation frequency, a single-T-DNA copy, the absence of a vector backbone, and stable and predictable transgene expression (2)?

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

I thank Drs. Walt Ream and Lan-Ying Lee for critical reading of this manuscript and for helpful suggestions regarding the text and figures, and I thank Dr. Seiichi Toki for permission to incorporate some of his models of T-DNA integration into **Figure 2**. Research in my laboratory is funded by the US National Science Foundation. I apologize to my colleagues whose work I did not cite because of space limitations.

LITERATURE CITED

1. Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, et al. 2003. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301:653–57
2. Altpeter F, Springer NM, Bartley LE, Blechl AE, Brutnell TP, et al. 2016. Advancing crop transformation in the era of genome editing. *Plant Cell* 28:1510–20
3. An S, Park S, Jeong D-H, Lee D-Y, Kang H-G, et al. 2003. Generation and analysis of end sequence database for T-DNA tagging lines in rice. *Plant Physiol.* 133:2040–47

4. Anand A, Krichevsky A, Schomack S, Lahaye T, Tzfira T, et al. 2007. *Arabidopsis* VIRE2 INTERACTING PROTEIN2 is required for *Agrobacterium* T-DNA integration in plants. *Plant Cell* 19:1695–708
5. Anand A, Vaghchhipawala Z, Ryu C-M, Kang L, Wang K, et al. 2007. Identification and characterization of plant genes involved in *Agrobacterium*-mediated plant transformation by virus-induced gene silencing. *Mol. Plant-Microbe Interact.* 20:41–52
6. Bakó L, Umeda M, Tiburcio AF, Schell J, Koncz C. 2003. The VirD2 pilot protein of *Agrobacterium*-transferred DNA interacts with the TATA box-binding protein and a nuclear protein kinase in plants. *PNAS* 100:10108–13
7. Bocharadt A, Hodal L, Palmgren G, Mattsson O, Okkels FT. 1992. DNA methylation is involved in maintenance of an unusual expression pattern of an introduced gene. *Plant Physiol.* 99:409–14
8. Brandle JE, McHugh SG, James L, Labbé H, Miki BL. 1995. Instability of transgene expression in field grown tobacco carrying the *csr1-1* gene for sulfonylurea herbicide resistance. *Bio/Technology* 13:994–98
9. Breyne P, Gheysen G, Jacobs A, Van Montagu M, Depicker A. 1992. Effect of T-DNA configuration on transgene expression. *Mol. Gen. Genet.* 235:389–96
10. Broothaerts W, Mitchell HJ, Weir B, Kaines S, Smith LMA, et al. 2005. Gene transfer to plants by diverse species of bacteria. *Nature* 433:629–33
11. Brunaud V, Balzergue S, Dubreucq B, Aubourg S, Samson F, et al. 2002. T-DNA integration into the *Arabidopsis* genome depends on sequences of pre-insertion sites. *EMBO Rep.* 3:1152–57
12. Char SN, Unger-Wallace E, Frame B, Briggs SA, Main M, et al. 2015. Heritable site-specific mutagenesis using TALENs in maize. *Plant Biotechnol. J.* 13:1002–10
13. Chen S, Jin W, Wang M, Zhang F, Zhou J, et al. 2003. Distribution and characterization of over 1000 T-DNA tags in rice genome. *Plant J.* 36:105–13
14. Chilton M-D, Que Q. 2003. Targeted integration of T-DNA into the tobacco genome at double-stranded breaks: New insights on the mechanism of T-DNA integration. *Plant Physiol.* 133:956–65
15. Citovsky V, Zupan J, Warnick D, Zambryski P. 1992. Nuclear localization of *Agrobacterium* VirE2 protein in plant cells. *Science* 256:1802–5
16. **Clark KA, Krysan PJ. 2010. Chromosomal translocations are a common phenomenon in *Arabidopsis thaliana* T-DNA insertion lines. *Plant J.* 64:990–1001**
17. Cluster PD, O'Dell M, Metzclaff M, Flavell RB. 1996. Details of T-DNA structural organization from a transgenic *Petunia* population exhibiting co-suppression. *Plant Mol. Biol.* 32:1197–203
18. **Crane YM, Gelvin SB. 2007. RNAi-mediated gene silencing reveals involvement of *Arabidopsis* chromatin-related genes in *Agrobacterium*-mediated root transformation. *PNAS* 104:15156–61**
19. Day CD, Lee E, Kobayashi J, Holappa LD, Albert H, Ow DW. 2000. Transgene integration into the same chromosome location can produce alleles that express at a predictable level, or alleles that are differentially silenced. *Genes Dev.* 14:2869–80
20. De Buck S, Jacobs A, Van Montagu M, Depicker A. 1998. *Agrobacterium tumefaciens* transformation and cotransformation frequencies of *Arabidopsis thaliana* root explants and tobacco protoplasts. *Mol. Plant-Microbe Interact.* 11:449–57
21. De Buck S, Jacobs A, Van Montagu M, Depicker A. 1999. The DNA sequences of T-DNA junctions suggest that complex T-DNA loci are formed by a recombination process resembling T-DNA integration. *Plant J.* 20:295–304
22. De Neve M, De Buck S, Jacobs A, Van Montagu M, Depicker A. 1997. T-DNA integration patterns in co-transformed plant cells suggest that T-DNA repeats originate from co-integration of separate T-DNAs. *Plant J.* 11:15–29
23. De Vos G, Zambryski P. 1989. Expression of *Agrobacterium* nopaline-specific VirD1, VirD2, and VirC1 proteins and their requirement for T-strand production in *E. coli*. *Mol. Plant-Microbe Interact.* 2:43–52
24. Deriano L, Roth DB. 2013. Modernizing the nonhomologous end-joining repertoire: Alternative and classical NHEJ share the stage. *Annu. Rev. Genet.* 47:433–55
25. Domínguez A, Fagoaga C, Navarro L, Moreno P, Peña L. 2002. Regeneration of transgenic citrus plants under non selective conditions results in high-frequency recovery of plants with silenced transgenes. *Mol. Genet. Genom.* 267:544–56
26. Dürrenberger F, Cramer A, Hohn B, Koukoliková-Nicola Z. 1989. Covalently bound VirD2 protein of *Agrobacterium tumefaciens* protects the T-DNA from exonucleolytic degradation. *PNAS* 86:9154–58

16. This paper describes the observation that chromosomal translocations are common following *Agrobacterium*-mediated transformation.

18. This paper describes several plant chromatin proteins important for efficient T-DNA integration.

27. Elmayan T, Vaucheret H. 1996. Expression of single copies of a strongly expressed 35S transgene can be silenced post-transcriptionally. *Plant J.* 9:787–97
28. Endo M, Ishikawa Y, Osakabe K, Nakayama S, Kaya H, et al. 2006. Increased frequency of homologous recombination and T-DNA integration in *Arabidopsis* CAF-1 mutants. *EMBO J.* 25:5579–90
29. Feng Z, Mao Y, Xu N, Zhang B, Wei P, et al. 2014. Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in *Arabidopsis*. *PNAS* 111:4632–37
30. Frame BR, Shou H, Chikwamba RK, Zhang Z, Xiang C, et al. 2002. *Agrobacterium tumefaciens*-mediated transformation of maize embryos using a standard binary vector system. *Plant Physiol.* 129:13–22
31. Francis KE, Spiker S. 2005. Identification of *Arabidopsis thaliana* transformants without selection reveals a high occurrence of silenced T-DNA integrations. *Plant J.* 41:464–77
32. Friesner J, Britt AB. 2003. *Ku80*- and *DNA ligase IV*-deficient plants are sensitive to ionizing radiation and defective in T-DNA integration. *Plant J.* 34:427–40
33. Gaj T, Gersbach CA, Barbas CF. 2013. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* 31:397–405
34. Gallego ME, Bleuyard J-Y, Daoudal-Cotterell S, Jallut N, White CI. 2003. Ku80 plays a role in non-homologous recombination but is not required for T-DNA integration in *Arabidopsis*. *Plant J.* 35:557–65
35. Gelvin SB. 2010. Finding a way to the nucleus. *Curr. Opin. Microbiol.* 13:53–58
36. Gelvin SB. 2010. Plant proteins involved in *Agrobacterium*-mediated genetic transformation. *Annu. Rev. Phytopathol.* 48:45–68
37. Gelvin SB. 2012. Traversing the cell: *Agrobacterium* T-DNA's journey to the host genome. *Front. Plant Sci.* 3:52
38. Gelvin SB, Kim S-I. 2007. Effect of chromatin upon *Agrobacterium* T-DNA integration and transgene expression. *Biochim. Biophys. Acta* 1769:410–21
39. Gheysen G, Villarroel R, Van Montagu M. 1991. Illegitimate recombination in plants: a model for T-DNA integration. *Genes Dev.* 5:287–97
40. Gohlke J, Deeken R. 2014. Plant responses to *Agrobacterium tumefaciens* and crown gall development. *Front. Plant Sci.* 5:155
41. Hamilton CM. 1997. A binary-BAC system for plant transformation with high-molecular-weight DNA. *Gene* 200:107–16
42. Hanin M, Volrath S, Bogucki A, Briker M, Ward E, Paszkowski J. 2001. Gene targeting in *Arabidopsis*. *Plant J.* 28:671–77
43. Harpster MH, Townsend JA, Jones JDG, Bedbrook J, Dunsmuir P. 1988. Relative strengths of the 35S cauliflower mosaic virus, 1', 2', and nopaline synthase promoters in transformed tobacco sugarbeet and oilseed rape callus tissue. *Mol. Gen. Genet.* 212:182–90
44. Henneberry TJ, Forlow Jech L, de la Torre T, Faulconer S, Hill JJ. 2000. *Pink bollworm egg infestations and larval survival in NuCOTN 33b and Deltapine cottons in Arizona*. Coll. Agric. Rep. AZ1170, Univ. Ariz., Tucson, AZ
45. Herrera-Estrella A, Chen Z-m, Van Montagu M, Wang K. 1988. 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.* 7:4055–62
46. Herrera-Estrella A, Van Montagu M, Wang K. 1990. A bacterial peptide acting as a plant nuclear targeting signal: The amino-terminal portion of *Agrobacterium* VirD2 protein directs a β -galactosidase fusion protein into tobacco nuclei. *PNAS* 87:9534–37
47. Hobbs SLA, Kpodar P, DeLong CMO. 1990. The effect of T-DNA copy number, position and methylation on reporter gene expression in tobacco transformants. *Plant Mol. Biol.* 15:851–64
48. Hodges LD, Cuperus J, Ream W. 2004. *Agrobacterium rhizogenes* GALLS protein substitutes for *Agrobacterium tumefaciens* single-stranded DNA-binding protein VirE2. *J. Bacteriol.* 186:3065–77
49. Hooykaas PJJ, Klapwijk PM, Nuti MP, Schilperoort RA, Rörsch A. 1977. Transfer of the *Agrobacterium tumefaciens* TI plasmid to avirulent *Agrobacteria* and to *Rhizobium ex planta*. *J. Gen. Microbiol.* 98:477–84
50. Howard E, Citovsky V. 1990. The emerging structure of the *Agrobacterium* T-DNA transfer complex. *BioEssays* 12:103–8
51. Hsing Y-I, Chern C-G, Fan M-J, Lu P-C, Chen K-T, et al. 2007. A rice gene activation/knockout mutant resource for high throughput functional genomics. *Plant Mol. Biol.* 63:351–64

52. Iglesias VA, Moscone EA, Papp I, Neuhuber F, Michalowski S, et al. 1997. Molecular and cytogenetic analyses of stably and unstably expressed transgene loci in tobacco. *Plant Cell* 9:1251–64
53. Iwakawa H, Carter BC, Bishop BC, Ogas J, Gelvin SB. 2017. Perturbation of H3K27me₃-associated epigenetic processes increases *Agrobacterium*-mediated transformation. *Mol. Plant-Microbe Interact.* 30:35–44
54. Jeong D-H, An S, Kang H-G, Moon S, Han J-J, et al. 2002. T-DNA insertional mutagenesis for activation tagging in rice. *Plant Physiol.* 130:1636–44
55. Jeong D-H, An S, Park S, Kang H-G, Park G-G, et al. 2006. Generation of a flanking sequence-tag database for activation-tagging lines in japonica rice. *Plant J.* 45:123–32
56. Jia Q, Bundock P, Hooykaas PJJ, de Pater S. 2012. *Agrobacterium tumefaciens* T-DNA integration and gene targeting in *Arabidopsis thaliana* non-homologous end-joining mutants. *J. Bot.* 2012:989272
57. Kilby NJ, Leyser HMO, Furner IJ. 1992. Promoter methylation and progressive transgene inactivation in *Arabidopsis*. *Plant Mol. Biol.* 20:103–12
58. Kim S-I, Veena, Gelvin SB. 2007. Genome-wide analysis of *Agrobacterium* T-DNA integration sites in the *Arabidopsis* genome generated under non-selective conditions. *Plant J.* 51:779–91
59. Kleinboelting N, Huep G, Appelhagen I, Viehoveer P, Li Y, Weisshaar B. 2015. The structural features of thousands of T-DNA insertion sites are consistent with a double-strand break repair-based insertion mechanism. *Mol. Plant* 8:1651–64
60. Köhler F, Cardon G, Pöhlman M, Gill R, Schieder O. 1989. 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.* 12:189–99
61. Kononov ME, Bassuner B, Gelvin SB. 1997. Integration of T-DNA binary vector ‘backbone’ sequences into the tobacco genome: evidence for multiple complex patterns of integration. *Plant J.* 11:945–57
62. Lacroix B, Citovsky V. 2016. A functional bacterium-to-plant DNA transfer machinery of *Rhizobium etli*. *PLOS Pathog.* 12(3):e1005502
63. Lacroix B, Loyter A, Citovsky V. 2008. Association of the *Agrobacterium* T-DNA–protein complex with plant nucleosomes. *PNAS* 105:15429–34
64. Lacroix B, Tzfira T, Vainstein A, Citovsky V. 2006. A case of promiscuity: *Agrobacterium*’s endless hunt for new partners. *Trends Genet.* 22:29–37
65. Lechtenberg B, Schubert D, Forsbach A, Gils M, Schmidt R. 2003. Neither inverted repeat T-DNA configurations nor arrangements of tandemly repeated transgenes are sufficient to trigger transgene silencing. *Plant J.* 34:507–17
66. Lee L-Y, Wu F-H, Hsu C-T, Shen S-C, Yeh H-Y, et al. 2012. Screening a cDNA library for protein–protein interactions directly in planta. *Plant Cell* 24:1746–59
67. Levy AA. 2016. T-DNA integration: Pol θ controls T-DNA integration. *Nat. Plants* 2:16170
68. Li J, Vaidya M, White C, Vainstein A, Citovsky V, Tzfira T. 2005. Involvement of Ku80 in T-DNA integration in plant cells. *PNAS* 102:19231–36
69. Li J-F, Norville JE, Aach J, McCormack M, Zhang D, et al. 2013. Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.* 31:688–91
70. Li J-F, Zhang D, Sheen J. 2014. Cas9-based genome editing in *Arabidopsis* and tobacco. *Methods Enzymol.* 546:459–72
71. Liang Z, Tzfira T. 2013. *In vivo* formation of double-stranded T-DNA molecules by T-strand priming. *Nat. Commun.* 4:2253
72. Liu L, Fan X-D. 2014. CRISPR–Cas system: a powerful tool for genome engineering. *Plant Mol. Biol.* 85:209–18
73. Loyter A, Rosenbluh J, Zakai N, Li J, Kozlovsky SV, et al. 2005. The plant VirE2 interacting protein 1. A molecular link between the *Agrobacterium* T-complex and the host cell chromatin? *Plant Physiol.* 138:1318–21
74. Magori S, Citovsky V. 2012. The role of the ubiquitin-proteasome system in *Agrobacterium tumefaciens*-mediated genetic transformation of plants. *Plant Physiol.* 160:65–71
75. Mahfouz MM, Piatek A, Stewart CN. 2014. Genome engineering via TALENs and CRISPR/Cas9 systems: challenges and perspectives. *Plant Biotechnol. J.* 12:1006–14

58. This paper describes the observation that, under nonselective conditions, T-DNA integrates randomly into the plant genome.

59. This paper details the rearrangements and origin of filler DNA at the junctions of numerous T-DNA insertions.

79. This paper shows that T-DNA integration can occur even when four different DNA repair/recombination pathways are simultaneously inactivated.

94. This paper shows that mutation of genes involved in NHEJ can result in increased T-DNA integration into the plant genome.

76. Majhi BB, Shah JM, Veluthambi K. 2014. A novel T-DNA integration in rice involving two interchromosomal translocations. *Plant Cell Rep.* 33:929–44
77. Martineau B, Voelker TA, Sanders RA. 1994. On defining T-DNA. *Plant Cell* 6:1032–33
78. Mayerhofer R, Koncz-Kalman Z, Nawrath C, Bakkeren G, Cramer A, et al. 1991. T-DNA integration: a mode of illegitimate recombination in plants. *EMBO J.* 10:697–704
79. **Mestiri I, Norre F, Gallego ME, White CI. 2014. Multiple host-cell recombination pathways act in *Agrobacterium*-mediated transformation of plant cells. *Plant J.* 77:511–20**
80. Meza TJ, Stangeland B, Mercy IS, Skårn M, Nymoen DA, et al. 2002. Analyses of single-copy *Arabidopsis* T-DNA-transformed lines show that the presence of vector backbone sequences, short inverted repeats and DNA methylation is not sufficient or necessary for the induction of transgene silencing. *Nucleic Acids Res.* 30:4556–66
81. Miranda A, Janssen G, Hodges L, Peralta EG, Ream W. 1992. *Agrobacterium tumefaciens* transfers extremely long T-DNAs by a unidirectional mechanism. *J. Bacteriol.* 174:2288–97
82. Mladenov E, Iliakis G. 2011. Induction and repair of DNA double strand breaks: the increasing spectrum of non-homologous end joining pathways. *Mutat. Res.* 711:61–72
83. Mlynárová L, Keizer LCP, Stiekema WJ, Nap J-P. 1996. Approaching the lower limits of transgene variability. *Plant Cell* 8:1589–99
84. Mlynárová L, Loonen A, Heldens J, Jansen RC, Keizer P, et al. 1994. Reduced position effect in mature transgenic plants conferred by the chicken lysozyme matrix-associated region. *Plant Cell* 6:417–26
85. Mysore KS, Bassuner B, Deng X, Darbinian NS, Motchoulski A, et al. 1998. Role of the *Agrobacterium tumefaciens* VirD2 protein in T-DNA transfer and integration. *Mol. Plant-Microbe Interact.* 11:668–683
86. Mysore KS, Nam J, Gelvin SB. 2000. An *Arabidopsis* histone H2A mutant is deficient in *Agrobacterium* T-DNA integration. *PNAS* 97:948–53
87. Nam J, Matthyse AG, Gelvin SB. 1997. Differences in susceptibility of *Arabidopsis* ecotypes to crown gall disease may result from a deficiency in T-DNA integration. *Plant Cell* 9:317–33
88. Narasimhulu SB, Deng X, Sarria R, Gelvin SB. 1996. Early transcription of *Agrobacterium* T-DNA genes in tobacco and maize. *Plant Cell* 8:873–86
89. Ni M, Cui D, Einstein J, Narasimhulu S, Vergara CE, Gelvin SB. 1995. Strength and tissue specificity of chimeric promoters derived from the octopine and mannopine synthase genes. *Plant J.* 7:661–76
90. Nishizawa-Yokoi A, Nonaka S, Saika H, Kwon Y-I, Osakabe K, Toki S. 2012. Suppression of Ku70/80 or Lig4 leads to decreased stable transformation and enhanced homologous recombination in rice. *New Phytol.* 196:1048–59
91. Oltmanns H, Frame B, Lee L-Y, Johnson S, Li B, et al. 2010. Generation of backbone-free, low transgene copy plants by launching T-DNA from the *Agrobacterium* chromosome. *Plant Physiol.* 152:1158–66
92. Ordon J, Gantner J, Kemna J, Schwalgun L, Reschke M, et al. 2017. Generation of chromosomal deletions in dicotyledonous plants employing a user-friendly genome editing toolkit. *Plant J.* 89:155–68
93. Pansegrau W, Schoumacher F, Hohn B, Lanka E. 1993. Site-specific cleavage and joining of single-stranded DNA by VirD2 protein of *Agrobacterium tumefaciens* Ti plasmids: analogy to bacterial conjugation. *PNAS* 90:11538–42
94. **Park S-Y, Vaghchhipawala Z, Vasudevan B, Lee L-Y, Shen Y, et al. 2015. *Agrobacterium* T-DNA integration into the plant genome can occur without the activity of key non-homologous end-joining proteins. *Plant J.* 81:934–46**
95. Peach C, Velten J. 1991. Transgene expression variability (position effect) of CAT and GUS reporter genes driven by linked divergent T-DNA promoters. *Plant Mol. Biol.* 17:49–60
96. Puchta H, Fauser F. 2014. Synthetic nucleases for genome engineering in plants: prospects for a bright future. *Plant J.* 78:727–41
97. Rossi L, Hohn B, Tinland B. 1996. Integration of complete transferred DNA units is dependent on the activity of virulence E2 protein of *Agrobacterium tumefaciens*. *PNAS* 93:126–30
98. Ruprecht C, Carroll A, Persson S. 2014. T-DNA-induced chromosomal translocation in *feronia* and *anxur2* mutants reveal implications for the mechanism of collapsed pollen due to chromosomal rearrangements. *Mol. Plant* 7:1591–94
99. Saika H, Nishizawa-Yokoi A, Toki S. 2014. The non-homologous end-joining pathway is involved in stable transformation in rice. *Front. Plant Sci.* 5:560

100. Sallaud C, Gay C, Larmande P, Bès M, Piffanelli P, et al. 2004. High throughput T-DNA insertion mutagenesis in rice: a first step towards *in silico* reverse genetics. *Plant J.* 39:450–64
101. Salomon S, Puchta H. 1998. Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *EMBO J.* 17:6086–95
102. Samson F, Brunaud V, Balzergue S, Dubreucq B, Lepiniec L, et al. 2002. FLAGdb/FST: a database of mapped flanking insertion sites (FSTs) of *Arabidopsis thaliana* T-DNA transformants. *Nucleic Acids Res.* 30:94–97
103. 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. *Plant J.* 80:1139–50
104. Schubert D, Lechtenberg B, Forsbach A, Gils M, Bahadur S, Schmidt R. 2004. Silencing in *Arabidopsis* T-DNA transformants: the predominant role of a gene-specific RNA sensing mechanism versus position effects. *Plant Cell* 16:2561–72
105. Shan Q, Wang Y, Li J, Gao C. 2014. Genome editing in rice and wheat using the CRISPR/Cas system. *Nat. Protoc.* 9:2395–410
106. Shi Y, Lee L-Y, Gelvin SB. 2014. Is VIP1 important for *Agrobacterium*-mediated transformation? *Plant J.* 79:848–60
107. Shurvinton CE, Hodges L, Ream W. 1992. A nuclear localization signal and the C-terminal omega sequence in the *Agrobacterium tumefaciens* VirD2 endonuclease are important for tumor formation. *PNAS* 89:11837–41
108. Simpson RB, O'Hara PJ, Kwok W, Montoya AL, Lichtenstein C, et al. 1982. DNA from the A6S/2 crown gall tumor contains scrambled Ti-plasmid sequences near its junctions with plant DNA. *Cell* 29:1005–14
109. Singer K, Shibolet Y, Li J, Tzfira T. 2012. Formation of complex extrachromosomal T-DNA structures in *Agrobacterium tumefaciens*-infected plants. *Plant Physiol.* 160:511–22
110. Sivamani E, Li X, Nalapalli S, Barron Y, Prairie A, et al. 2015. Strategies to improve low copy transgenic events in *Agrobacterium*-mediated transformation of maize. *Transgen. Res.* 24:1017–27
111. Song J, Bent AF. 2014. Microbial pathogens trigger host DNA double-strand breaks whose abundance is reduced by plant defense responses. *PLOS Pathog.* 10(4):e1004030
112. Stachel SE, Messens E, Van Montagu M, Zambryski P. 1985. Identification of the signal molecules produced by wounded plant cells that activate T-DNA transfer in *Agrobacterium tumefaciens*. *Nature* 318:624–29
113. Stachel SE, Nester EW. 1986. The genetic and transcriptional organization of the *vir* region of the A6 Ti plasmid of *Agrobacterium tumefaciens*. *EMBO J.* 5:1445–54
114. Stachel SE, Timmerman B, Zambryski P. 1987. Activation of *Agrobacterium tumefaciens vir* gene expression generates multiple single-stranded T-strand molecules from the pTiA6 T-region: requirement for 5' *virD* gene products. *EMBO J.* 6:857–63
115. Strizhov N, Li Y, Rosso MG, Viehoveer P, Dekker KA, Weisshaar B. 2003. High-throughput generation of sequence indexes from T-DNA mutagenized *Arabidopsis thaliana* lines. *BioTechniques* 35:1164–68
116. Szabados L, Kovács I, Oberschall A, Abrahám E, Kerekes I, et al. 2002. Distribution of 1000 sequenced T-DNA tags in the *Arabidopsis* genome. *Plant J.* 32:233–42
117. Tax FE, Vernon DM. 2001. T-DNA-associated duplication/translocations in *Arabidopsis*. Implications for mutant analysis and functional genomics. *Plant Physiol.* 126:1527–38
118. Tenea GN, Spantzel J, Lee L-Y, Zhu Y, Lin K, et al. 2009. Overexpression of several *Arabidopsis* histone genes increases *Agrobacterium*-mediated transformation and transgene expression in plants. *Plant Cell* 21:3350–67
119. Thole V, Worland B, Wright J, Bevan MW, Vain P. 2010. Distribution and characterization of more than 1000 T-DNA tags in the genome of *Brachypodium distachyon* community standard line Bd21. *Plant Biotechnol. J.* 8:734–47
120. Tinland B, Hohn B, Puchta H. 1994. *Agrobacterium tumefaciens* transfers single-stranded transferred DNA (T-DNA) into the plant cell nucleus. *PNAS* 91:8000–4
121. Tzfira T, Frankman LR, Vaidya M, Citovsky V. 2003. Site-specific integration of *Agrobacterium tumefaciens* T-DNA via double-stranded intermediates. *Plant Physiol.* 133:1011–23

101. This paper shows that T-DNA can be captured into induced double-strand DNA breaks.

109. This paper shows that T-DNA circles can form in planta.

118. This paper shows that overexpression of specific histone proteins can enhance stable *Agrobacterium*-mediated transformation.

130. This paper shows that DNA polymerase θ is required for T-DNA integration into the *Arabidopsis* genome.

122. Tzfira T, Li J, Lacroix B, Citovsky V. 2004. *Agrobacterium* T-DNA integration: molecules and models. *Trends Genet.* 20:375–83
123. Tzfira T, Vaidya M, Citovsky V. 2001. VIP1, an *Arabidopsis* protein that interacts with *Agrobacterium* VirE2, is involved in VirE2 nuclear import and *Agrobacterium* infectivity. *EMBO J.* 20:3596–607
124. Tzfira T, Vaidya M, Citovsky V. 2004. Involvement of targeted proteolysis in plant genetic transformation by *Agrobacterium*. *Nature* 431:87–92
125. Tzfira T, Weinthal D, Marton I, Zeevi V, Zuker A, Vainstein A. 2012. Genome modifications in plant cells by custom-made restriction enzymes. *Plant Biotechnol. J.* 10:373–89
126. Ülker B, Li Y, Rosso MG, Logemann E, Somssich IE, Weisshaar B. 2008. T-DNA-mediated transfer of *Agrobacterium tumefaciens* chromosomal DNA into plants. *Nat. Biotechnol.* 26:1015–17
127. Vaghchhipawala ZE, Vasudevan B, Lee S, Morsy MR, Mysore KS. 2012. *Agrobacterium* may delay plant nonhomologous end-joining DNA repair via XRCC4 to favor T-DNA integration. *Plant Cell* 24:4110–23
128. van Attikum H, Bundock P, Hooykaas PJJ. 2001. Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. *EMBO J.* 20:6550–58
129. van Attikum H, Bundock P, Overmeer RM, Lee L-Y, Gelvin SB, Hooykaas PJJ. 2003. The *Arabidopsis* *AtLIG4* gene is required for the repair of DNA damage, but not for the integration of *Agrobacterium* T-DNA. *Nucleic Acids Res.* 31:4247–55
130. van Kregten M, de Pater S, Romeijn R, van Schendel R, Hooykaas PJJ, Tijsterman M. 2016. T-DNA integration in plants results from polymerase- θ -mediated DNA repair. *Nat. Plants* 2:16164
131. Veluthambi K, Ream W, Gelvin SB. 1988. Virulence genes, borders, and overdrive generate single-stranded T-DNA molecules from the A6 Ti plasmid of *Agrobacterium tumefaciens*. *J. Bacteriol.* 170:1523–32
132. Vergunst AC, Hooykaas PJJ. 1998. Cre/lox-mediated site-specific integration of *Agrobacterium* T-DNA in *Arabidopsis thaliana* by transient expression of cre. *Plant Mol. Biol.* 38:393–406
133. Vergunst AC, Jansen LET, Hooykaas PJJ. 1998. Site-specific integration of *Agrobacterium* T-DNA in *Arabidopsis thaliana* mediated by Cre recombinase. *Nucleic Acids Res.* 26:2729–34
134. Ward ER, Barnes WM. 1988. VirD2 protein of *Agrobacterium tumefaciens* very tightly linked to the 5' end of T-strand DNA. *Science* 242:927–30
135. Waterworth WM, Masnavi G, Bhardwaj RM, Jiang Q, Bray CM, West CE. 2010. A plant DNA ligase is an important determinant of seed longevity. *Plant J.* 63:848–60
136. Wei F-J, Kuang L-Y, Oung H-M, Cheng S-Y, Wu H-P, et al. 2016. Somaclonal variation does not preclude the use of rice transformants for genetic screening. *Plant J.* 85:648–59
137. Weising K, Bohn H, Kahl G. 1990. Chromatin structure of transferred genes in transgenic plants. *Dev. Genet.* 11:233–47
138. Wenck A, Czakó M, Kanevski I, Márton L. 1997. Frequent collinear long transfer of DNA inclusive of the whole binary vector during *Agrobacterium*-mediated transformation. *Plant Mol. Biol.* 34:913–22
139. Wendt T, Doohan F, Winckelmann D, Mullins E. 2011. Gene transfer into *Solanum tuberosum* via *Rhizobium* spp. *Transgenic Res.* 20:377–86
140. Winans SC. 1992. Two-way chemical signaling in *Agrobacterium*-plant interactions. *Microbiol. Rev.* 56:12–31
141. Windels P, De Buck S, Van Bockstaele E, De Loose M, Depicker A. 2003. T-DNA integration in *Arabidopsis* chromosomes. Presence and origin of filler DNA sequences. *Plant Physiol.* 133:2061–68
142. Wolterink-van Loo S, Escamilla Ayala AA, Hooykaas PJJ, van Heudsdén GPH. 2015. Interaction of the *Agrobacterium tumefaciens* virulence protein VirD2 with histones. *Microbiology* 161:401–10
143. Yadav NS, Vanderleyden J, Bennett DR, Barnes WM, Chilton M-D. 1982. Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *PNAS* 79:6322–26
144. Yusibov VM, Steck TR, Gupta V, Gelvin SB. 1994. Association of single-stranded transferred DNA from *Agrobacterium tumefaciens* with tobacco cells. *PNAS* 91:2994–98
145. Zaltsman A, Krichevsky A, Kozlovsky SV, Yasmin F, Citovsky V. 2010. Plant defense pathways subverted by *Agrobacterium* for genetic transformation. *Plant Signal. Behav.* 5:1245–48
146. Zaltsman A, Lacroix B, Gafni Y, Citovsky V. 2013. Disassembly of synthetic *Agrobacterium* T-DNA-protein complexes via the host SCF^{VB} ubiquitin-ligase complex pathway. *PNAS* 110:169–74

147. Zhi L, TeRonde S, Meyer S, Arling ML, Register JC, et al. 2015. Effect of *Agrobacterium* strain and plasmid copy number on transformation frequency, event quality and usable event quality in an elite maize cultivar. *Plant Cell Rep.* 34:745–54
 148. Zhu Y, Nam J, Humara JM, Mysore KS, Lee L-Y, et al. 2003. Identification of *Arabidopsis* *rat* mutants. *Plant Physiol.* 132:494–505
 149. Ziemienowicz A, Tinland B, Bryant J, Gloeckler V, Hohn B. 2000. Plant enzymes but not *Agrobacterium* VirD2 mediate T-DNA ligation in vitro. *Mol. Cell. Biol.* 20:6317–22
-

RELATED RESOURCES

- Tzfira T, Hohn B, Gelvin SB. 2017. Transfer of genetic information from *Agrobacterium* to plants. In *Reference Module in Life Sciences*, ed. SD O'Neill. San Diego: Elsevier. <https://doi.org/10.1016/B978-0-12-809633-8.07278-2>
- Wang K, ed. 2015. *Agrobacterium Protocols*, Vol. 2. Methods Mol. Biol. Ser. New York: Humana Press. <https://link.springer.com/book/10.1007%2F978-1-4939-1658-0>
- Williams ME, Yuan Z. 2012. A really useful pathogen, *Agrobacterium tumefaciens*. *Plant Cell* 24: tpc.112.tt1012