Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 6169–6173, September 1987
Botany

Integration of *Agrobacterium tumefaciens* transfer DNA (T-DNA) involves rearrangements of target plant DNA sequences

(*Nicotiana tabacum*)

GODELIEVE GHEYSEN, MARC VAN MONTAGU*, AND PATRICIA ZAMBRYSKI†

Laboratorium voor Genetica, Rijksuniversiteit Gent, B-9000 Gent, Belgium

Contributed by Marc Van Montagu, May 26, 1987

ABSTRACT **The transfer DNA (T-DNA) mobilized into plant cells by *Agrobacterium tumefaciens* seems to integrate rather randomly into the plant genome. We analyzed a target site in the genome of *Nicotiana tabacum* before and after integration of a T-DNA. Clones presenting right and left T-DNA/plant DNA junctions were used as probes to identify and isolate a unique 1.8-kilobase *Eco*RI fragment corresponding to the plant DNA target site for a T-DNA insertion event. Comparison of the nucleotide sequences of the plant DNA portions of the T-DNA junction clones with the original plant DNA target revealed that several types of rearrangements resulted from insertion of the T-DNA. The most dramatic alteration was a 158-base-pair direct repeat of target plant sequences at the left and right T-DNA junctions. In addition, there were deletion and insertion events at the ends of the right and left copies of the 158-base-pair repeat. The variety of target-site rearrangements suggests that T-DNA insertion is a multistep process of recombination accompanied by local replicative and repair activities mediated by host-cell enzymes.**

Agrobacterium tumefaciens induces crown gall tumors on wounded plants (for a review, see ref. 1). During this transformation, a specific segment of the bacterial Ti plasmid, the transfer DNA (T-DNA), is transferred across bacterial and plant cell membranes and becomes stably integrated into the plant nuclear genome. The tumorous phenotype of the transformed cells is caused by the action of several T-DNA genes. It has been shown that none of these or other T-DNA genes are needed to transfer the T-DNA to the plant cell or integrate it into the nuclear genome (2). The only structural sequences needed for T-DNA transfer are the 25-base-pair (bp) direct repeats that flank the T-DNA on the Ti plasmid. These repeats are cis-acting sequences that direct T-DNA transfer in a polar right-to-left fashion (3, 4); any DNA that is located between these T-DNA borders is efficiently transferred and integrated into the plant cell.

The functional products that mediate T-DNA transfer are encoded outside the T-DNA region by the Ti plasmid virulence (*vir*) loci (5) and the chromosomal virulence (*chv*) loci (6). While the *chv* genes are constitutively expressed, the *vir* genes are organized as a single regulon whose expression is activated by specific phenolic compounds present in exudates of wounded cells (7). During the first hours after induction of *vir* gene expression, molecular reactions associated with T-DNA transfer begin to occur within *Agrobacterium*. A linear single-stranded free T-DNA homologous molecule, designated the T-strand, is produced (8). This molecule corresponds to the bottom strand of the T-DNA region so that its 5′ and 3′ ends map to the right and left border repeats, respectively. *vir* induction also results in the production of site-specific cleavages of the 25-bp border repeats within the bottom strand on the Ti plasmid (8–10). Border nicks have been proposed to represent initiation (right border) and termination (left border) sites for T-strand synthesis (8). The specific properties of the T-strand molecule have led to the formation of a model for the mechanism of T-DNA transfer to plant cells analogous to bacterial conjugation (8, 11).

While some of the early bacterial reactions that lead to T-DNA transfer are beginning to be understood, little is known about later events that lead to T-DNA integration into the plant nuclear genome. To date, analyses have been limited to sequence determination of the junctions between T-DNA ends and plant DNA (12–14). These studies reveal that junction points at the right end are localized close to or in the 25-bp repeat. The junction points at the left end are spread over about 100 bp, suggesting that the T-DNA does not show the precision of transposons in integration. However, the most important deficiency in this analysis is that the plant DNA target sites utilized during integration have never been characterized. In this paper, we present data on the nucleotide sequence of a plant DNA target prior to and after T-DNA integration.

MATERIALS AND METHODS

Bacterial Strains and Plasmids. *Escherichia coli* KH802 (su^- derivative of K802; ref. 15) was used for growing phage λ Ch35, and *E. coli* MC1061 (16) was used for plasmid transformation. Plasmids cg3 and cg4 are subclones of right and left T-DNA/plant DNA junction fragments, respectively; the cloning and characterization of these fragments and the tumor cell line of their origin, *Nicotiana tabacum* W38C58-1, have been described (12, 17).

DNA Analysis. Restriction endonucleases and phage T4 DNA ligase were purchased from Boehringer Mannheim, Bethesda Research Laboratories (N.V. GIBCO, Belgium), and New England Biolabs. DNA·DNA hybridizations (12) and DNA sequencing (18) were performed as described. Unless specified, all other methods have been described (19).

Cloning of Plant DNA. Nuclear DNA was prepared from leaves of *N. tabacum* W38 as described (20), digested with *Eco*RI, and fractionated on a sucrose gradient. Plant DNA in the size range of 1.2–2.5 kilobases (kb) was ligated to purified Ch35 *Eco*RI arms; the ligation reaction contained 1.4 μg of plant DNA and 0.7 μg of Ch35 arms in a 10-μl volume. Ligated DNA was packaged with Gigapack (Stratagene, Genofit, Switzerland), yielding 2×10^6 phages per μg of DNA. Recombinant phages were screened on Bio-Assay dishes (Nunc; GIBCO, Belgium) using nylon filters (Amersham). Recombinant phages contained an average of seven

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: T-DNA, transfer DNA.

*To whom reprint requests should be addressed.
†Present address: Division of Molecular Plant Biology, University of California, Berkeley, CA 94720.

*Eco*RI fragments, and approximately 1 plaque in 10^4 hybridized to the cg3 probe.

RESULTS

Identification and Cloning of a T-DNA Target Sequence. To isolate a T-DNA target site in plant DNA requires the identification of an integration event that can be purified by molecular cloning techniques; only T-DNAs integrated adjacent to nonrepetitive DNA can be used for target DNA cloning. The W38C58-1 tumor line of *N. tabacum* contains a suitable T-DNA integration event. Southern blot hybridization indicated several T-DNA copies in this tumor line (17). One right (cg3) T-DNA/plant DNA junction fragment and one left one (cg4) were cloned from total W38C58-1 tumor cell DNA digested with *Eco*RI (12). cg3 contains T-DNA sequences up to the first nucleotide of the right copy of the 25-bp border repeat linked to ≈550 bp of tobacco DNA. cg4 contains ≈1300 bp of tobacco DNA linked to T-DNA sequences starting 92 nucleotides inside the left copy of the 25-bp repeat. The plant DNAs representing the cg3 and cg4 integration events were characterized by using cg3 and cg4 DNAs as hybridization probes to Southern blots of *Eco*RI-digested total DNA from untransformed tobacco cells. While cg4 hybridized to repeated sequences distributed throughout the entire molecular weight range of the *Eco*RI-digested DNA, cg3 hybridized to a single 1.8-kb *Eco*RI fragment (12). Therefore, cg3 is suitable as a hybridization probe to isolate a plant DNA target sequence for a T-DNA integration event.

Total untransformed tobacco DNA was digested with *Eco*RI, and the fraction enriched for DNAs of 1.2–2.5 kb was cloned into the phage λ vector Ch35 (21). Recombinant phages were screened with cg3 as probe, and 10 "target" candidates were isolated. Because *N. tabacum* W38 is an allodiploid, isolated target sequences may not be identical. However, all 10 candidates contained 1.8-kb *Eco*RI fragments with identical restriction patterns upon further digestion with *Hae* III. The subclone of a 1.8-kb *Eco*RI fragment into pBR322 was named pTT1 (T-DNA target 1).

Although the pTT1 target was identified by its homology to the unique plant DNA in cg3, it also hybridized to repetitive sequences in tobacco DNA (data not shown); and the pattern of multiple bands hybridizing to pTT1 resembled that found when cg4 was utilized as a probe against total tobacco DNA. Moreover, the sum of the plant DNA sequences in cg3 (0.55 kb) and cg4 (1.3 kb) closely matched that of pTT1 (1.8 kb). Therefore, we reasoned that cg3 and cg4 might be derived from the same T-DNA insertion. Indeed, restriction mapping and hybridization analyses confirmed that cg4 was homologous to pTT1 (see below). Thus, pTT1, cg3, and cg4 contain the plant DNA sequences corresponding to a particular plant target before and after T-DNA integration (Fig. 1).

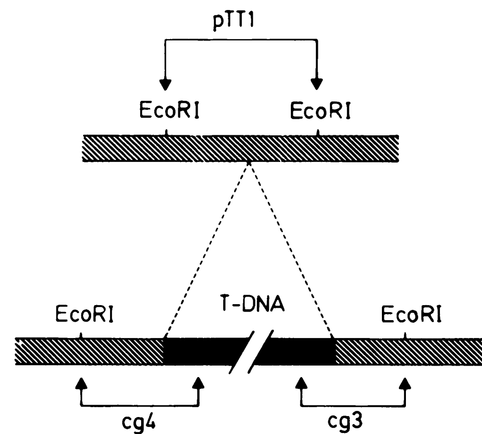

FIG. 1. Schematic representation of the isolated clones. (*Upper*) Diagram showing the unoccupied target site. (*Lower*) The corresponding site after T-DNA integration. ▨, Plant DNA; ■, T-DNA; pTT1, target; cg4, left border junction; cg3, right border junction. The *Eco*RI sites correspond to those in Fig. 2.

Restriction Mapping of Two T-DNA Border Clones and the Corresponding Target Sequence. Fig. 2 shows the *Hae* III maps for cg4, pTT1, and cg3; the dashed vertical lines align the regions of homology between cg4 and pTT1 and between cg3 and pTT1. The 470-bp *Hae* III fragment of pTT1 was deduced to contain the target site for T-DNA integration because it was the only fragment not present in either cg3 or cg4. The T-DNA/plant DNA junction fragments of cg3 and cg4 were similarly deduced to be 750-bp and 1050-bp *Hae* III

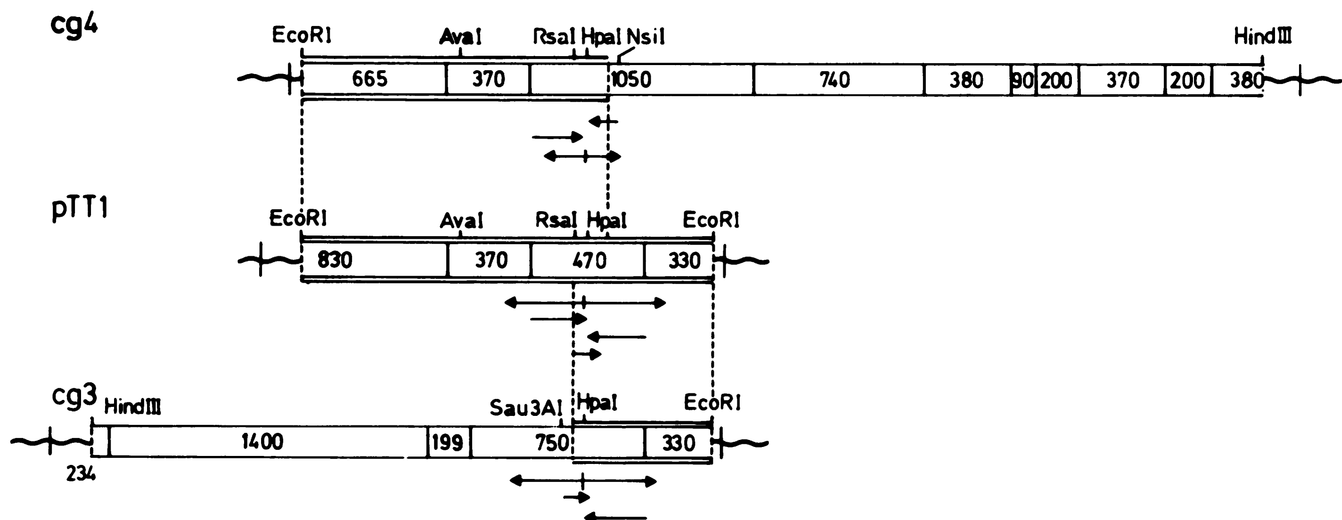

FIG. 2. *Hae* III restriction map of pTT1, cg3, and cg4. Single-lined boxes, T-DNA; double-lined boxes, plant DNA; wavy lines, pBR322. Dashed vertical lines align the corresponding plant sequences in the different clones; cg3 and cg4 are *Hind*III–*Eco*RI fragments in pBR322 (12), whereas pTT1 is an *Eco*RI clone in pBR322. Vertical lines indicate *Hae* III restriction sites; numbers refer to the sizes of the *Hae* III fragments. The plant DNAs in the 665-bp (cg4) and 830-bp (pTT1) *Hae* III fragments are identical (about 650 bp); the sizes of the fragments differ because the inserts were cloned in opposite orientations into pBR322 (the distance from the left *Eco*RI site to the closest *Hae* III site in pBR322 is 16 bp in cg4 and 177 bp in pTT1). Some other restriction enzymes used in the analyses are indicated. The direction and length of the horizontal arrows summarize the strategy used to obtain the nucleotide sequences shown in Fig. 3.

fragments, respectively [i.e., the sizes of these fragments differ from pTT1 and from the known right or left T-DNA-border *Hae* III fragments (12)]. A sequencing strategy was defined to compare the exact points of divergence of the different DNAs (Fig. 2).

The *Hae* III digests and additional mapping data suggested that no major rearrangements had occurred in the plant DNA of the junction clones compared to the plant target DNA. Heteroduplex analyses confirmed that the plant DNA portions of cg3 and cg4 form heteroduplex complexes with pTT1 DNA in the absence of visible loop structures (data not shown).

Nucleotide Sequence Comparison of the Plant DNA Target Before and After T-DNA Integration. The sequences presented in Fig. 3 reveal that the right-end junction site of cg3 is entirely contiguous with T-DNA or plant DNA sequences (i.e., there is a direct transition from T-DNA to plant DNA sequences at a site corresponding to position 201 of the pTT1 sequence). Surprisingly, there was a deletion of 27 bp in the plant DNA portion of cg3 at a distance of 165 bp from the junction point (Fig. 3 *Middle*, positions 366–392). The origin of this deletion and the accompanying C → T transition (Fig. 3 *Middle*, position 393) are better understood after comparison with the sequences at the left T-DNA border in cg4 (see below).

The transition from the left end of the T-DNA to plant DNA is not direct. The nucleotide sequence at the junction of cg4 reveals the presence of an unexpected sequence of 33 bp (Fig. 3, *Top*, positions 359–391) which occurs between the T-DNA and the plant DNA sequences of pTT1. This "filler" DNA sequence is homologous to an imperfect, inverted repeat at positions 562–589 in pTT1 and to a short direct repeat at positions 307–317 in pTT1 (Fig. 3, arrows).

The most dramatic rearrangement of target DNA sequences was the duplication of a 158-bp segment from pTT1; these 158 bp are present at both the left and right T-DNA borders (Fig. 3, boxed sequences, corresponding to positions 201–358 of pTT1). The generation of this duplication is likely related to the formation of the deletion at positions 366–392 of cg3 and the "filler" DNA at the junction point of cg4, since these latter events occurred at the ends of the duplicated segment. These data together suggest that T-DNA integration leads to several types of rearrangements of plant DNA close to the site of the T-DNA insertion.

Characteristics of the T-DNA Target Sequence. T-DNA integration can occur at numerous chromosomal locations in the plant genome (22, 23); however, at the nucleotide level integration may require or prefer specific sequences. While the sequence of pTT1 is the only target site available, its base composition close to the integration site is noteworthy. The average A+T content of *N. tabacum* is about 60% (24), and the A+T content of pTT1 is close to this value at positions 1–106 (58%) and 228–595 (67%). However, the site where the right end of the T-DNA invaded the plant DNA (position 200) is flanked by 94 nucleotides on the left and 28 nucleotides on the right that are extremely A+T-rich (91.7% and 78.6%, respectively). The plant DNA of other published T-DNA junctions is also often A+T-rich (12), and high A+T content has been seen in a variety of other integration targets (e.g., *Alu* transposition; ref. 25). Such sequences may result in local destabilization of duplex plant DNA and thereby enhance opportunities for an integration event.

The plant DNAs of pTT1 and cg4 contain repetitive sequences. To localize these repeats more precisely, purified subfragments of pTT1 and cg4 were hybridized to W38 plant DNA digested with *Eco*RI. Sequences to the right of the *Ava* I site (Fig. 2) in pTT1 or cg4 hybridized to a single *Eco*RI fragment, and sequences to the left of this site (>500 bp from the T-DNA insertion site) hybridized to multiple fragments (data not shown). Previous studies have shown that at least

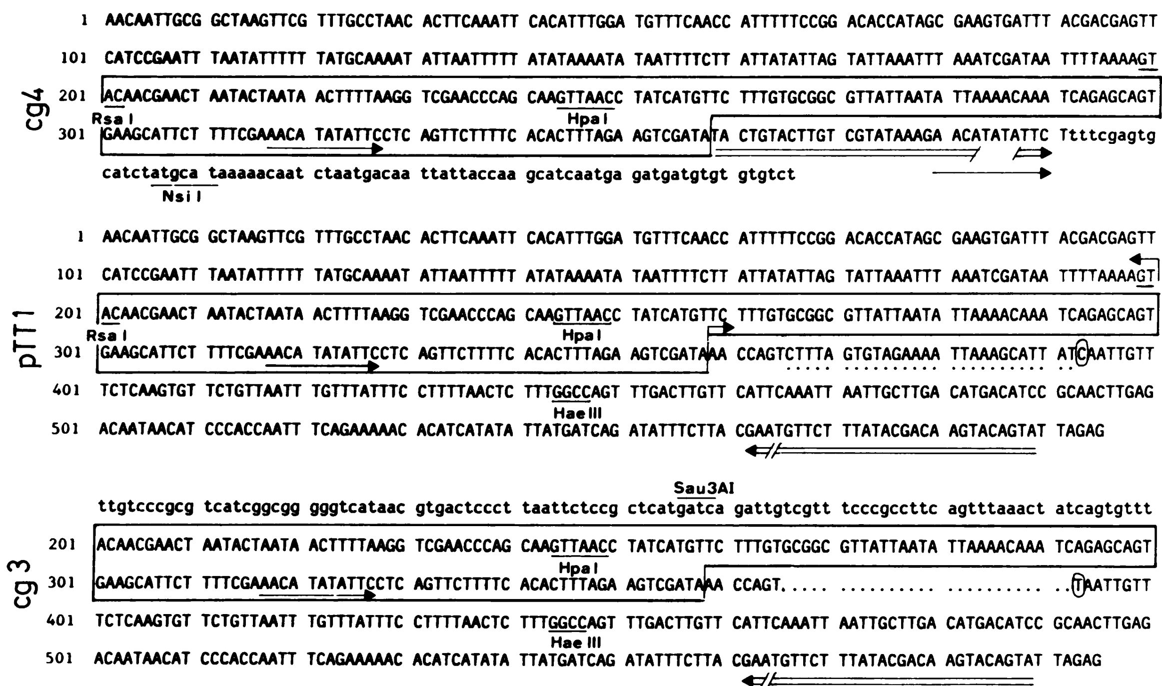

FIG. 3. Nucleotide sequence of a plant DNA site before and after T-DNA integration. Large and small characters indicate plant DNA and T-DNA, respectively. (*Top*) cg4 (left T-DNA junction). (*Middle*) pTT1 sequence. (*Bottom*) cg3 (right T-DNA junction). Only plant DNA (not T-DNA) sequences are numbered, and numbers in cg3 and cg4 refer to sequences identical to pTT1. Bent arrows in pTT1 indicate points where T-DNA/plant DNA junctions of cg3 (position 200) and cg4 (position 358) occur. Boxed sequences indicate a 158-bp region of pTT1 (positions 201–358) that occurs at both cg3 and cg4 junctions. In cg3, the deletion (positions 366–392) is indicated by dots and the C → T transition (position 393) is circled. In cg4, possible origins for the "filler" sequence (positions 359–391) are indicated by arrows in pTT1 (because these sequences are also present in cg3 and cg4, they are marked there as well). Single-lined arrows mark an 11-bp homology, and double-lined arrows mark an interrupted (25 bp + 3 bp) inverted homology.

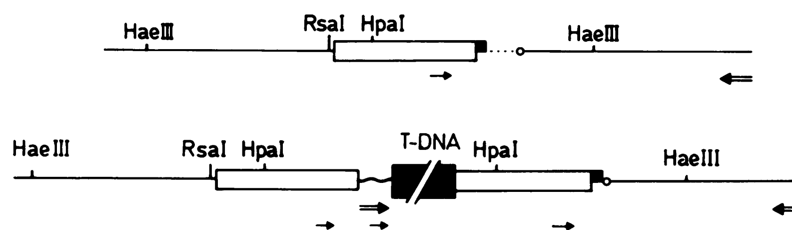

FIG. 4. Summary of the rearrangements made by a T-DNA insertion. (*Upper*) Diagram of the unoccupied target site; the *Hae* III sites delineate the 470-bp fragment (Fig. 2). (*Lower*) Diagram of the left and right border junctions of the T-DNA insertion. The T-DNA is drawn as an interrupted black box, and the 158-bp target duplication is shown as an empty box. On the right we find the 158-bp duplication (empty box), an unchanged 7-bp sequence (black square, corresponding to positions 359–365 of Fig. 3), the deletion of target sequences (dots), and the C → T transition (circle, corresponding to position 393 of Fig. 3). On the left, rearrangements include the 158-bp duplication and "filler" DNA (zigzag line). Arrows indicate homology of "filler" DNA to other sequences. Restriction enzyme sites in the plant DNA correspond to those in Figs. 2 and 3. For details, see the text and Fig. 3.

55% of the DNA of *N. tabacum* is organized as single-copy sequences, on average 1400 bp long, alternating with repetitive sequences ≈300 bp long (24). These data together imply that, even if the T-DNA frequently integrates close to or within repetitive plant sequences, plant target sites appropriate for cloning may be identified by testing subfragments of T-DNA/plant DNA junction clones for their ability to hybridize to unique plant DNA.

DISCUSSION

We show that plant DNA sequences can undergo rearrangements as a result of integration of the T-DNA. Fig. 4 summarizes the plant DNA target analyzed here before and after T-DNA integration. The most striking alteration of target sequences is the generation of a direct repeat of 158 bp at the left and right T-DNA junctions. Besides this duplication event, there are additional sequence alterations, such as a deletion and insertion at the ends of the 158-bp segment. These data indicate that T-DNA integration is unlikely to occur in a single step.

There are two general mechanisms for DNA insertion into the genome of an eukaryotic cell. Either the DNA element itself encodes some or most of the enzymes involved in integration (we refer to such systems as type I events, exemplified by transposons and retroviruses), or DNA insertion is mediated by host enzymes normally used for DNA synthesis, recombination, or repair; such integrations (type II events) are observed upon introduction of DNA through microinjection, electroporation, or transfection, and upon infection with certain viruses, such as polyoma, simian virus 40, or adenovirus. Table 1 summarizes the characteristics of both systems and compares their features with those known for T-DNA integration.

Type I systems promote integration in a relatively undisturbing manner; usually a short segment (<10 bp) of target sequences is duplicated at the ends of the insertion element, and further rearrangements are rare. In contrast, such a large duplication as seen here following T-DNA integration has been observed after translocation of cellular oncogenes during the development of malignant plasmacytomas (30). These latter events are postulated to arise by imprecise recombination promoted by cellular enzymes.

T-DNA integration also resembles type II systems in the variety of rearrangements that occur at the target site. For example, deletion of target sequences is typical for type II events, and their extent varies from a few (31) to several thousand (28) nucleotides. In addition, the insertion of extra "filler" nucleotides during the formation of recombinant joints occurs in many type II events. At a polyoma virus/host joint, a 37-bp "filler" sequence is an inverted perfect duplication of a single-copy host sequence found 650 bp away from the virus/host junction (28). Similarly, the 33-bp "filler" sequence at the plant DNA/T-DNA junction in cg4 is unique and homologous to an inverted host sequence about 200 bp away from the junction. The probability that this homology is fortuitous is approximately 2×10^{-10}. The finding of "filler" sequences in both animal and plant cell DNA may indicate that the (repair) mechanism responsible for their generation occurs throughout most eukaryotic cells. Perhaps the junction sequences that occur between tandemly arranged T-DNA copies (12, 13) also represent "filler" sequences generated by a similar mechanism.

There are no obvious homologies between sequences at the ends of the T-DNA and cellular sequences immediately adjacent to the junction sites. However, there is a plant DNA sequence, T-G-C-G-G-C-G-T-T-A-T-T-A-A-T (Fig. 3, positions 265–279) that is partially homologous to a sequence,

Table 1. Comparison of different integration systems

General characteristics of integration event	Integration systems*		
	Type I	Type II	T-DNA
Integration	Efficient	Usually inefficient	Efficient
DNA configuration	Always single units	Mostly tandem arrays	Singly or in (short) tandem arrays
Integrity of DNA	Unaltered structure	Often modified	Occasionally modified
Target duplication	Small (<10 bp); element specific	None or variable	Large (158-bp) duplication†
Deletions of target DNA	Uncommon	General feature	27-bp deletion†
Filler DNA at junctions	Uncommon	Frequent	33-bp filler†
Enzymes used	Both element- and host-encoded enzymes	Host enzymes‡	Unknown†

*Type I, transposons and retroviruses (26); type II, simian virus 40, polyoma, and artificially introduced DNA (27, 28); T-DNA (refs. 1, 12, and 13 and this work).
†The only available data (this work), therefore not to be generalized.
‡Although polyoma large T antigen can influence the efficiency and mode of viral integration, it has been shown not to be essential for the integration process itself (29).

T-G-C-G-G-A-C-G-T-T-T-T-A-A-T, located 31 bp internal to the left 25-bp T-DNA border repeat (12). These sequences may have promoted an initial contact between target plant sequences and invading T-DNA. An 11-bp potential contact sequence near an octopine-type left T-DNA junction site also has been observed (13). Contact sequences have been proposed to facilitate polyoma (32) or pseudogene (33) integration as well.

T-DNA integration is most different from type II events in that it is a highly efficient process. Only microinjection of DNA into the nucleus shows a comparable transformation efficiency (34). This might indicate that the T-DNA is directed preferentially into the nucleus and/or is possibly protected by a protein complex from nucleolytic degradation.

The right T-DNA/plant DNA junction analyzed here contains a direct transition from T-DNA to plant DNA sequences; in contrast, the left T-DNA junction contains scrambled "filler" sequences. The precision of the right junction compared to the left may be explained if T-DNA integration initiates with the right T-DNA end. This proposal is consistent with the observed functional polarity of the T-DNA border sequences (3, 4). Polar T-DNA transfer and/or integration could potentially be mediated by a "pilot" protein linked to the right end of the T-DNA during its transit (3).

It has been proposed recently that the T-DNA transfer intermediate is a single-stranded copy of the T-DNA region, the T-strand. While it is unknown whether the T-strand is altered (e.g., into a duplex form) prior to transfer and/or integration, for the present discussion we assume the T-strand is the transfer intermediate. That the T-strand is single-stranded is not inconsistent with its role as an integration intermediate, since it has been shown (35) that single-stranded DNA can be used to transfect mammalian cells. Furthermore, models for general recombination involve invasion of target sequences by single-stranded donor DNA (36).

While it is still too early to formulate a precise mechanism for T-DNA insertion into plant DNA, the following steps are consistent with the available information. First, the T-strand is transferred to the plant cell as a DNA-protein complex. Second, a protein at the 5′ (right) end of the T-strand interacts with a nicked sequence in plant DNA. Third, local torsional strain on the plant DNA results in the production of a second nick on the opposite strand of the target sequence. The position of the second nick may vary with different T-DNA insertion events. Fourth, the T-strand is ligated to plant DNA, and its homologous strand is copied by cellular enzymes. Fifth, repair and replication of the staggered nick in the plant target results in the production of a repeated sequence (of variable length) and additional sequence rearrangements ("filler" DNA, deletion, etc.) at the ends of the inserted T-DNA element. Although the T-DNA is a specialized system for the integration of bacterial DNA in the plant genome, we do not find clear evidence for the involvement of *Agrobacterium*-encoded enzymes in integration. The rearrangements observed here at the plant DNA/T-DNA junctions rather indicate a role for plant-encoded recombination/repair activities.

The genetic transformation of plant cells by *Agrobacterium* is a fundamentally interesting process that is also used widely as a tool to genetically engineer plant cells with DNAs of interest. The present studies provide information on the events underlying the insertion of the T-DNA into the plant genome and help to predict more specific applications of the T-DNA transfer system. The T-DNA has already proven to be an efficient vector for the transfer of intact sequences to the plant. However, the rearrangements observed here in the plant target sequences advise a certain caution in experiments that rely on precise fusion between target plant DNA and invading T-DNA sequences. Analysis of additional plant DNA targets should reveal whether the rearrangements observed here are a general result of T-DNA integration.

The authors thank Jan Gielen for help with sequence analysis; Gilbert Engler for heteroduplex analysis; Allan Caplan, Peter Czernilofsky, Ann Depicker, and Liz Howard for critical reading of the manuscript; and Martine De Cock, Janice Quartieri, Karel Spruyt, and Stefaan Van Gijsegem for preparation of the manuscript and figures. This work was supported by grants of the Algemene Spaar-en Lijfrentekas-Kankerfonds, the Fonds voor Wetenschappelijk Geneeskundig Onderzoek (F.G.W.O. 3.0001.82), and the Services of the Prime Minister (12.0561.84) to M.V.M. and J. Schell; G.G. is grateful to the National Fund for Scientific Research (Belgium) for a former Research Assistant fellowship.

1. Gheysen, G., Dhaese, P., Van Montagu, M. & Schell, J. (1985) in *Genetic Flux in Plants*, Advances in Plant Gene Research, eds. Hohn, B. & Dennis, E. S. (Springer, Vienna), Vol. 2, pp. 11–47.
2. Leemans, J., Deblaere, R., Willmitzer, L., De Greve, H., Hernalsteens, J.-P., Van Montagu, M. & Schell, J. (1982) *EMBO J.* **1**, 147–152.
3. Wang, K., Herrera-Estrella, L., Van Montagu, M. & Zambryski, P. (1984) *Cell* **38**, 455–462.
4. Peralta, E. G. & Ream, L. W. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 5112–5116.
5. Stachel, S. E. & Nester, E. W. (1986) *EMBO J.* **5**, 1445–1454.
6. Douglas, C. J., Staneloni, R. J., Rubin, R. A. & Nester, E. W. (1985) *J. Bacteriol.* **161**, 850–860.
7. Stachel, S. E., Messens, E., Van Montagu, M. & Zambryski, P. (1985) *Nature (London)* **318**, 624–629.
8. Stachel, S. E., Timmerman, B. & Zambryski, P. (1986) *Nature (London)* **322**, 706–712.
9. Yanofsky, M. F., Porter, S. G., Young, C., Albright, L. M., Gordon, M. P. & Nester, E. W. (1986) *Cell* **47**, 471–477.
10. Wang, K., Stachel, S., Timmerman, B., Van Montagu, M. & Zambryski, P. (1987) *Science* **235**, 587–591.
11. Stachel, S. E. & Zambryski, P. (1986) *Cell* **47**, 155–157.
12. Zambryski, P., Depicker, A., Kruger, K. & Goodman, H. (1982) *J. Mol. Appl. Genet.* **1**, 361–370.
13. Holsters, M., Villarroel, R., Gielen, J., Seurinck, J., De Greve, H., Van Montagu, M. & Schell, J. (1983) *Mol. Gen. Genet.* **190**, 35–41.
14. Kwok, W. W., Nester, E. W. & Gordon, M. P. (1985) *Nucleic Acids Res.* **13**, 459–471.
15. Wood, W. B. (1966) *J. Mol. Biol.* **16**, 118–133.
16. Casadaban, M. J. & Cohen, S. N. (1980) *J. Mol. Biol.* **138**, 179–207.
17. Lemmers, M., De Beuckeleer, M., Holsters, M., Zambryski, P., Depicker, A., Hernalsteens, J.-P., Van Montagu, M. & Schell, J. (1980) *J. Mol. Biol.* **144**, 353–376.
18. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
19. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
20. Fischer, R. L. & Goldberg, R. B. (1982) *Cell* **29**, 651–660.
21. Loenen, W. A. M. & Blattner, F. R. (1983) *Gene* **26**, 171–179.
22. Ambros, P. F., Matzke, A. J. M. & Matzke, M. A. (1986) *EMBO J.* **5**, 2073–2077.
23. Chyi, Y.-S., Jorgensen, R. A., Goldstein, D., Tanksley, S. D. & Loaiza-Figueroa, F. (1986) *Mol. Gen. Genet.* **204**, 64–69.
24. Zimmerman, J. L. & Goldberg, R. B. (1977) *Chromosoma* **59**, 227–252.
25. Daniels, G. R. & Deininger, P. L. (1985) *Nucleic Acids Res.* **13**, 8939–8954.
26. Shapiro, J. A., ed. (1983) *Mobile Genetic Elements* (Academic, New York).
27. Weinberg, R. A. (1980) *Annu. Rev. Biochem.* **49**, 197–226.
28. Williams, T. J. & Fried, M. (1986) *Mol. Cell. Biol.* **6**, 2179–2184.
29. Della Valle, G., Fenton, R. G. & Basilico, C. (1981) *Cell* **23**, 347–355.
30. Gerondakis, S., Cory, S. & Adams, J. M. (1984) *Cell* **36**, 973–982.
31. Dejean, A., Bougueleret, L., Grzeschik, K.-H. & Tiollais, P. (1986) *Nature (London)* **322**, 70–72.
32. Ruley, H. E. & Fried, M. (1983) *Nature (London)* **304**, 181–184.
33. Van Arsdell, S. W. & Weiner, A. M. (1984) *Nucleic Acids Res.* **12**, 1463–1471.
34. Capecchi, M. R. (1980) *Cell* **22**, 479–488.
35. Rauth, S., Song, K.-Y., Ayares, D., Wallace, L., Moore, P. D. & Kucherlapati, R. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 5587–5591.
36. Whitehouse, H. L. K., ed. (1982) *Genetic Recombination* (Wiley, New York).