# Tumor DNA Structure in Plant Cells Transformed by *A. tumefaciens*

Patricia Zambryski, Marcelle Holsters, Kelly Kruger
Ann Depicker, Josef Schell, Marc Van Montagu
Howard M. Goodman

*Summary*. Crown gall tumors are induced in plants by infection with the soil bacterium *Agrobacterium tumefaciens*. Because the tumor induction involves transfer of a portion of the tumor-inducing (Ti) plasmid DNA from the bacterium to the plant cells, this system is of interest for the study of genetic exchange as well as tumor induction. The boundaries of the transferred DNA (T-DNA) have been cloned from transformed plant cells of tobacco. Detailed mapping with restriction enzymes and nucleotide sequence analysis of two independent clones were used to study the molecular structure of the ends of the T-DNA. One clone contains the two ends of the T-DNA joined together; the other contains one end of the T-DNA joined to repetitive plant DNA sequences. These studies provide direct evidence that the T-DNA can be integrated into the plant genome. In addition, the data suggest that in the plant, T-DNA can be tandemly repeated. Sequence analysis of the junction of crown gall clone 1 reveals several direct repeats as well as an inverted repeat; these structures may be involved in the transfer of the DNA from *Agrobacterium* to plant cells.

Crown gall is a neoplastic disease produced in most dicotyledonous plants by infection of the plant with the Gram-negative soil bacterium *Agrobacterium tumefaciens*. Crown gall tumors can proliferate autonomously in tissue culture and, in some cases, it is possible to regenerate plants from these cells (*1*). The tumorous state of crown gall tissue has been shown to result from the transfer of DNA from *A. tumefaciens* to the plant cell [see reviews cited in (*2*, *3*)]. This natural system for DNA exchange has begun to be studied extensively, since in addition to providing an interesting system of genetic exchange and tumor induction, it may make available a vector for the introduction of new genes into plants by molecular genetic engineering methods.

It is now firmly established by both physical and genetic evidence that the agent responsible for crown gall tumor induction is a large plasmid, Ti (tumor-inducing), contained in all oncogenic strains of *A. tumefaciens* (*4*). Loss of these large plasmids results in loss of bacterial oncogenicity, and introduction of a Ti plasmid into originally nononcogenic plasmid-free *Agrobacterium* strains renders them oncogenic.

Crown gall cells produce derivatives of basic amino acids (such as arginine or lysine) called opines (*5*). In nature, *Agrobacterium* utilizes opines as sources of carbon and nitrogen. The genetic information for the synthesis and catabolism of opines is carried by the Ti plasmid itself; hence Ti plasmids are classified according to the type of opine synthesized in tumor cells. *Agrobacterium* strains which induce the synthesis of nopaline [*N*-α(1,3-dicarboxypropyl)-L-arginine] in crown gall cells (nopaline strains) and strains which induce the synthesis of octopine [*N*-α(*N*-l-carboxyethyl)-L-arginine] (octopine strains) have been studied most extensively. Thus, the interaction of *Agrobacterium* with plants is a novel sort of parasitism, which has been described as genetic colonization (*2*). *Agrobacterium tumefaciens* introduces genetic information into its host and thereby induces the host to synthesize products coded for by the newly introduced genes, which it alone can utilize.

The Ti plasmids of tumor-inducing strains of *A. tumefaciens* are large, ranging in size from 150 to more than 200 kilobase pairs (kbp) (*6*-*9*), and encode functions for oncogenicity, opine biosynthesis and catabolism, conjugative transfer between bacteria, sensitivity to antibiotics, and exclusion of bacteriophage Ap1 (*10*). There are several regions of DNA which are common to both octopine and nopaline Ti plasmids; one of the common DNA sequences is contained within the DNA region which is present in transformed plant cells (*6*, *8*, *11*). This transferred DNA is called T-DNA. Kinetic (*12*) and blot hybridization (*2*, *13*, *14*) analyses have provided evidence that the T-DNA can stably integrate into the genome of transformed plant cells.

Almost nothing is known about the mechanism of transfer of Ti plasmid DNA to plants. Most likely, the transfer involves two stages: first, the bacterial and plant cell walls interact in a manner which allows the Ti plasmid to enter the plant cell [the bacteria themselves do not enter (*15*)], and second, part of the Ti plasmid DNA is transferred to plant DNA.

The work described in this article is concerned with the second stage of the transfer. We analyzed the DNA in a tumor cell line of tobacco, the T37 teratoma isolated by Braun (*1*, *16*), which was induced by the nopaline Ti plasmid strain T37. Hybridization blotting analyses of the plant DNA isolated from the T37 teratoma and other nopaline tumor cell lines (*14*) suggest that the insertion process responsible for the T-DNA transfer probably recognizes specific sequences at the ends of the T-DNA; the length of the T-DNA segment found in association with the plant DNA is relatively constant. The T-DNA of independent tumors induced by nopaline Ti plasmid has been shown to be a continuous stretch of DNA 23 kbp in length, colinear with the T region of the Ti plasmid (*14*).

Figure 1 shows possible models for the transfer of T-DNA to plant DNA. Either the T-DNA interacts directly with plant DNA, or the ends of the T-DNA first interact with each other prior to insertion into plant DNA. To investigate this pro-

The authors' affiliations are: P. Zambryski, K. Kruger, and H. M. Goodman, Department of Biochemistry and Biophysics and Howard Hughes Medical Institute, University of California, San Francisco (UCSF) 94143; M. Holsters and A. Depicker, Department of Biochemistry and Biophysics, UCSF, and Laboratorium voor Genetika en Histologie, Rijksuniversiteit Gent, B-9000 Gent, Belgium; J. Schell, Laboratorium voor Genetika en Histologie, Rijksuniversiteit Gent, and Max-Planck-Institut für Zuchtungsforschung, Köln, D-5000 West Germany; and M. Van Montagu, Laboratorium voor Genetika en Histologie, Rijksuniversiteit Gent. Correspondence should be sent to Dr. Goodman.


0036-8075/80/0919-1385$01.75/0  

cess, we isolated the boundaries of the T-DNA from the T37 teratoma by molecular cloning techniques and studied their structure.

### Molecular Cloning of the Boundaries of T-DNA from Transformed Plant Cells

An Eco RI and Hind III restriction map of the T region of nopaline strain C58 Ti plasmid (*9*) is shown in Fig. 2A. A different natural isolate, nopaline strain T37, was used to induce the teratoma tumor cell line of tobacco (*Nicotiana tabacum* cv. Havana) used in this study (the restriction map of the T region of C58 and T37 is compared below). The T-DNA of transformed plant cells from the T37 teratoma has been mapped in detail (*14*). The T-DNA is colinear with the Ti plasmid, since all the internal Eco RI and Hind III fragments from Hind III fragment A to Hind III fragment B are found in tumor DNA (Fig. 2A). The ends of the T-DNA were defined by hybridization, using radioactive probes synthesized from specific restriction fragments isolated from this region of the Ti plasmid. On the left, the T-DNA ends in Hind III fragment B (6.3 kbp) between Eco RI fragments of 1.5 and 3.9 kbp; on the right, the T-DNA ends in Hind III fragment A (3.15 kbp) contained within the 14.1-kbp Eco RI fragment. Details of the determination of the T-DNA boundaries in teratoma T37 and other tumor cell lines will be published (*14*). In the present study, the T-DNA boundary fragments in plant tumor DNA are defined as those fragments which hybridize to probes of fragments A or B and which are different in size from the corresponding fragments in the Ti plasmid. When plant tumor DNA from T37 teratoma is digested with Eco RI there are two fragments of 8.7 and 12.4 kbp which fulfill these criteria. These two fragments were cloned in the bacteriophage lambda vector Charon 4A (*17*) and their structures determined by restriction mapping and hybridization analysis as diagrammed in Fig. 2B.

The restriction enzyme digestion pattern of crown gall (cg) clone 1 (12.4 kbp) and cg clone 2 (8.7 kbp) were compared to the pattern from the T region of Ti plasmid (Fig. 2B). Crown gall clone 2 is colinear with the Ti plasmid to a point within fragment A. However, cg clone 1 has an unexpected structure; it is colinear with the Ti plasmid to a point within fragment A on one side and is also colinear to a point within fragment B on the other side. (To illustrate its structure, we show cg clone 1 bracketing two tandem

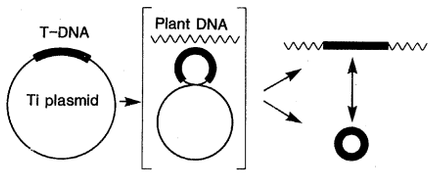



Fig. 1. Schematic representation of several possible models of transfer of T-DNA from Ti plasmid to plant cell DNA.

T-DNA regions in Fig. 2B.) Thus, cg clone 1 hybridizes to probes of fragments A and B and implies a joining of the two ends of the T-DNA. Crown gall clone 2 hybridizes to fragment A only and represents a single end of the T-DNA. We have not yet obtained a clone for the other end of the T-DNA with homology to fragment B.

The crosshatched areas in Fig. 2B are the border fragments of the T-DNA. The Hind III/Hind III fragment AcgB (5.2 kbp) from cg clone 1 and the Hind III/Eco RI fragment Acg (2.9 kbp) from cg clone 2 were subcloned into the appropriate sites in the plasmid pBR322 to facilitate further analysis.

### Restriction Enzyme Analysis of the Border Fragments of T-DNA

The two border DNA fragments Acg and AcgB were analyzed in detail to determine whether the T-DNA integrates into the plant cell genome during tumor induction by the Ti plasmid. Two experimental approaches were taken. First, a detailed restriction map of fragments Acg and AcgB was constructed to determine whether there were any subfragments not derived from either of the Ti plasmid fragments A or B. Second, fragments Acg and AcgB were used as probes in hybridization to total DNA from untransformed plants.

Five different DNA's were digested with endonuclease Hae III (pBR322, Ti plasmid fragments A and B in pBR322, and fragments Acg and AcgB in pBR322) and the resultant fragments were separated by electrophoresis (Fig. 3E). After transfer to nitrocellulose, the Hae III fragments from the gel separation were hybridized with specific radioactive probes [to purified fragments A (Fig. 3A), B (Fig. 3B), Acg (Fig. 3C), or AcgB (Fig. 3D)].

The results of these experiments (summarized in Fig. 3, F and G) indicate that the T-DNA boundary for Acg is contained within the Hae III fragment of 1200 base pairs (bp), and for AcgB the T-DNA boundary is within the Hae III fragment of 550 bp on the fragment A portion, while on the fragment B portion there are unexpectedly two Hae III fragments of 840 and 880 bp which hybridize to B but are different in size from B fragments.

For the analysis of Fig. 3, cloned fragments A and B were available from the nopaline plasmid C58 (*9*). By overall restriction pattern (*9*) and electron microscope heteroduplex analysis (*8*), the T region of nopaline plasmid C58 is identical to the T region of nopaline plasmid T37, except for a small deletion in the 4.2-kbp Eco RI fragment of C58 which reduces its size to 2.85 kbp in T37. Nevertheless, it seemed important, especially with respect to the observation of two different boundary "B-like fragments" in clone AcgB, to demonstrate conclusively that the observed differences in Hae III patterns between Ti plasmid fragments A and B and fragments AcgB and Acg were not due to differences between nopaline strains T37 and C58, since T37 was used to induce the plant tumor used in this study.

Fragments A and B from T37 Ti plasmid were therefore cloned in pBR322. The Hae III digestion patterns were compared (see Fig. 4 and below) and fragments A and B were found to be identical in the two nopaline strains, except that there is an 840-bp Hae III fragment in T37 fragment B instead of the 740-bp Hae III fragment of C58; that is, there is a 100-bp deletion in C58 between the Hinc II and the Hae III site of the 840-bp fragment of T37. This new Hae III fragment of 840 bp is identical in size to the 840-bp Hae III fragment in fragment AcgB of cg clone 1. Thus, these results explain the aforementioned discrepancy, and in AcgB there is only one Hae III fragment of 880 bp which hybridizes to B but is different in size than any B Hae III fragments.

The order of the Hae III fragments in each of the clones A, Acg, AcgB, and B was determined by restriction mapping and hybridization analysis (Fig. 4). We have included the Hae III map of fragment B from both T37 and C58 nopaline Ti plasmids for comparison. Additional restriction enzymes were used in the mapping and their sites are included in this diagram. The data presented in Figs. 3 and 4 can be summarized as follows:

1) There are no unique Hae III fragments of Acg or AcgB which do not hybridize to either fragment A or fragment B. Thus, there is no totally foreign DNA (for example, plant) contained in these clones which has at least two Hae III sites.

2) There is no Hae III fragment in AcgB which hybridizes to both frag-

ments A and B. (As will be shown below, the Hae III site occurs close to the junction of the A and B sequences.)

3) The A portion of Acg ends to the "right" of the Sac II site (see Fig. 4 and below). The A portion of AcgB also ends to the right of the Sac II site, at or very close to the Hae III junction; the 150-bp Sac II/Hae III fragment hybridizes uniquely to Ti plasmid A and not B.

4) The B portion of AcgB ends to the right of the Eco RI site of the 1240-bp Hae III fragment of B (note the Eco RI site present in B is missing in AcgB); no fragments to the left of this Eco RI site in the T37 "B" clone hybridize to an AcgB probe.

The Sac II/Eco RI DNA fragment of Acg and the Sac II/Kpn I DNA fragment of AcgB contain the T-DNA boundary sequences of interest, and these fragments were used in further studies. Figure 5 shows the Hinf I restriction enzyme maps for the Sac II/Kpn I fragment of AcgB (labeled fragment 1) and the Sac II/Eco RI fragments of Acg (labeled fragment 2); these maps are shown expanded from the Hae III restriction maps of AcgB and Acg (Fig. 4). Fragments 1 and 2 were digested with Hinf I and hybridized with probes made from fragments A and/or B of the Ti plasmid (Fig. 5, B and C). The Hinf I divides the Sac II/Eco RI fragment of Acg into a portion of 190 bp which hybridizes to fragment A and a portion of approximately 600 bp which does not hybridize to fragment A (Fig. 5C). In contrast, all the Hinf I fragments of the Sac II/Kpn I

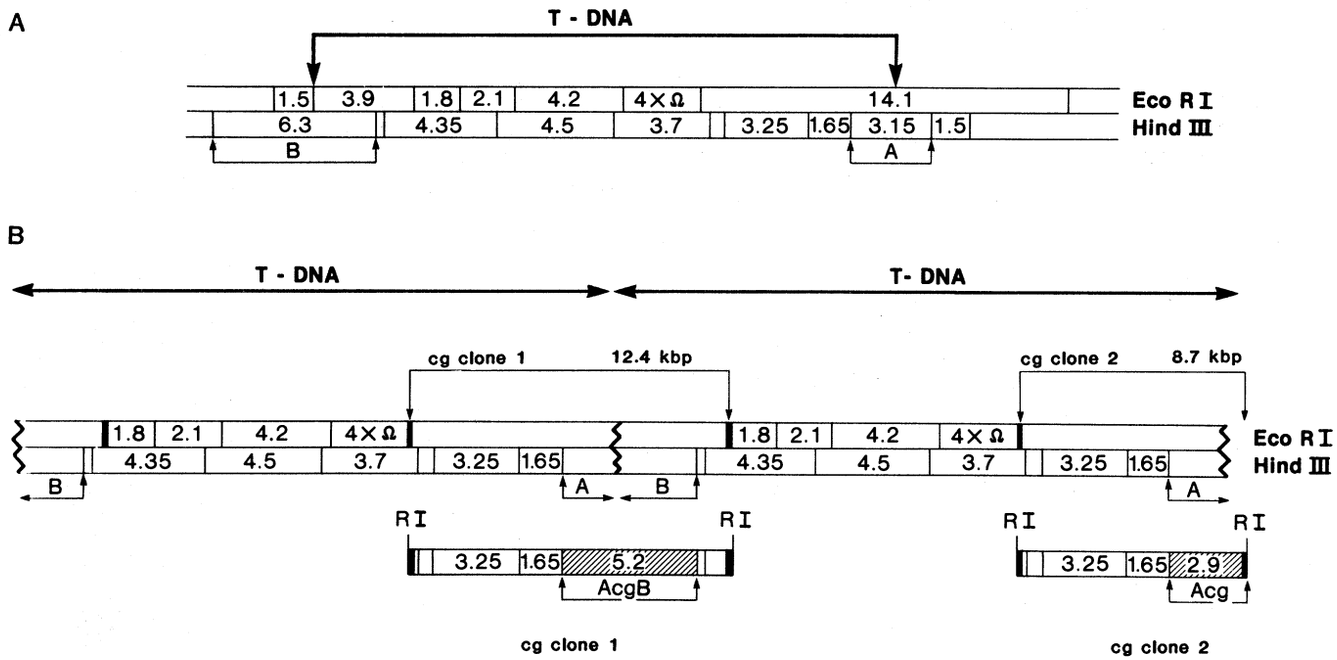
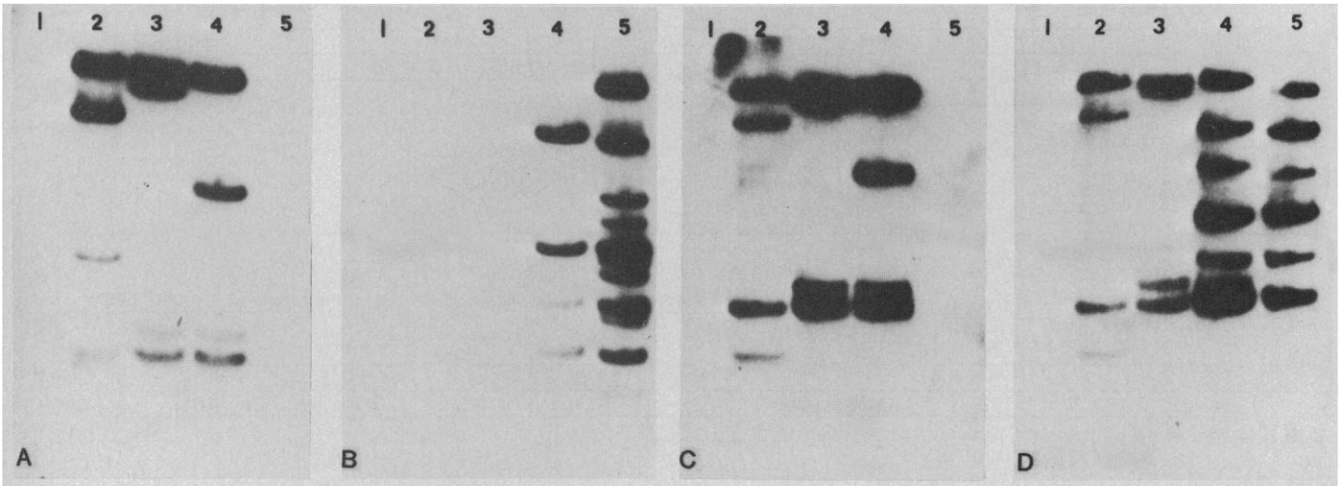



Fig. 2. The T-DNA of nopaline Ti plasmids. (A) Map of the T-DNA region of nopaline Ti plasmid C58 for the restriction enzymes Eco RI and Hind III, drawn to scale in kilobase pairs. The restriction map of Ti plasmid T37 is identical except for a small deletion in the 4.2-kbp Hind III fragment which reduces its size in T37 to 2.85 kbp (see also text and Figs. 3 and 4). The four small Eco RI fragments (4 × Ω) are 0.6, 0.7, 0.8, and 0.9 kbp, but their order has not been determined. The small Hind III fragment between the 6.3- and 4.35-kbp fragments is 0.3 kbp, and the one between the 3.7- and 3.25-kbp fragments is 0.6 kbp; their sizes are not shown on the map because of lack of space. Large arrows indicate T-DNA, the portion of the Ti plasmid found in plant tumor cells. Small arrows indicate Hind III fragments A (3.15 kbp) and B (6.3 kbp), which contain the borders of the T-DNA. These fragments were previously referred to as Hind III fragments 23 and 10, respectively (*9*). (B) Schematic representation of the boundaries of T-DNA in the T37 teratoma cell line. Two tandem copies of the T-DNA are drawn. The T-DNA junctions are shown as jagged lines in the upper part and as crosshatched areas below. The T-DNA junctions were cloned with the EK2 vector bacteriophage Charon 4A (*17*), yielding cg clone 1 and cg clone 2. Charon 4A and recombinant phage were propagated in *E. coli* strain DP50 supF (*17*) under P1 physical containment in accordance with NIH guidelines for recombinant DNA research. Phage DNA was purified as described in (*24*). Charon 4A "arms" were prepared by sucrose gradient fractionation after Eco RI digestion (*25*). For the isolation of cg clone 2, 6 μg of total teratoma DNA digested with Eco RI was ligated to 12 μg of "arms." For the isolation of cg clone 1, the teratoma DNA digested with Eco RI was enriched for DNA ranging from 10 to 15 kbp by sucrose gradient fractionation and 3 μg of this DNA was ligated to 12 μg of "arms." Vector DNA concentration in the ligation reaction (50 m*M* tris, *p*H 8, 7 m*M* $MgCl_2$, 1 m*M* dithiothreitol, 0.1 m*M* EDTA, 1 m*M* ATP plus T4 DNA ligase from Biolab at 20 U/ml at 4°C for 12 hours) was always 600 μg/ml. DNA was packaged into phage (*17*) and recombinant phage were screened [by the method of Benton and Davis (*26*) as modified by Cordell *et al.* (*24*)], using radioactive probes specific to fragment A or B of Ti plasmid. Crown gall clones 1 and 2 were digested with Hind III or Eco RI, and the restriction fragments were separated on 1 percent agarose gels and transferred to nitrocellulose (Schleicher and Schuell) (*27*). Fragments were tested for homology to radioactively labeled fragments of Ti plasmid. Hybridizations were carried out at 42°C in 50 percent formamide, 5 × strength Denhardt solution (*28*), triple-strength SSC (1 × SSC is 0.15*M* NaCl, 0.015*M* sodium citrate), 20 μg/ml *E. coli* carrier DNA, 40 μg/ml yeast carrier RNA, 50 m*M* Hepes (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid), *p*H 7.5. After hybridization, filters were washed in 0.1 × SSC containing 0.1 percent sodium dodecyl sulfate (SDS) at 50°C. The cg clone 1 hybridizes to both Ti plasmid fragments A and B; it is homologous to the Ti plasmid from the Eco RI site at the left edge of the Eco RI 14.1-kbp fragment to the Eco RI site between the Eco RI 3.9- and 1.8-kbp fragments. The Hind III fragment AcgB (crosshatching) contains the DNA junction between the two ends of the T-DNA in cg clone 1. The cg clone 2 contains a single boundary of the T-DNA; by hybridization it is homologous to the Ti plasmid from the Eco RI site at the left edge of the 14.1-kbp Eco RI fragment to a point just before a new Eco RI site. The Hind III/Eco RI fragment Acg (crosshatching) is homologous in part to Ti plasmid fragment A. The Eco RI restriction sites at the ends of cg clones 1 and 2 are indicated by heavy lines; the Hind III sites are indicated internal to the Eco RI sites in the lower diagram. The cg clone 1 DNA in λ Charon 4A was digested with Hind III and ligated to pBR322 DNA digested with Hind III and treated with calf intestinal phosphatase (Calbiochem), and the ligation mix was used to transform *E. coli* HB101 (*29*). Plasmids were screened for the presence of the 5.2-kbp fragment AcgB, which hybridizes to radioactive probes specific for fragment A or B. Similarly, cg clone 2 DNA was digested with both Hind III and Eco RI and ligated to pBR322 DNA digested with Hind III and Eco RI. After transformation of *E. coli* HB101, recombinant plasmids were screened for the presence of the 2.9-kbp fragment Acg, which hybridizes specifically to radioactive fragment A.

fragment of AcgB hybridize to either A or B specific probes (Fig. 5B). There is a Hinf I fragment of 165 bp which hybridizes weakly in Fig. 5B (shown as a cross-hatched band in the diagram in Fig. 5, A and B). Preliminary nucleotide sequence analysis reveals that this fragment has a composition of about 80 percent A-T (A, adenine; T, thymine) and is homologous to the nucleotide sequence of Ti plasmid fragment B in this region (only approximately 92 bp have been compared between B and AcgB in this region so far). Its base composition and small size may prevent this fragment from hybridizing efficiently to fragment B specific probe.

When fragment 1 is digested with Hinf I two fragments of 310 bp are produced. To separate these two fragments, fragment 1 was first digested with Hae III and the 880-bp Hae III fragment subsequently digested with Hinf I (the sizes of the digestion products are shown in Fig. 5A). The 235-bp Hae III/Hinf I fragment hybridizes to Ti plasmid fragment B specific probe (data not shown); however, by nucleotide sequence analysis it is not completely homologous to B, as indicated by the stippled region in Fig. 5A (see below).

The Sac II/Eco RI fragment of Acg and the Sac II/Kpn I fragment of AcgB were used as hybridization probes against total tobacco DNA digested with Eco RI. Under stringent hybridization conditions [50 percent formamide, 42°C, triple-strength SSC (see legend to Fig. 2)], we do not detect hybridization of these DNA's to plant DNA. However, under less stringent hybridization conditions (50 percent formamide, 25°C, triple-strength SSC) the Sac II/Eco RI fragment of Acg hybridizes to the entire molecular weight range of Eco RI fragments of total plant DNA, while the Sac II/Kpn I fragment of AcgB does not (Fig. 5D). These data suggest that the plant DNA in Acg is representative of



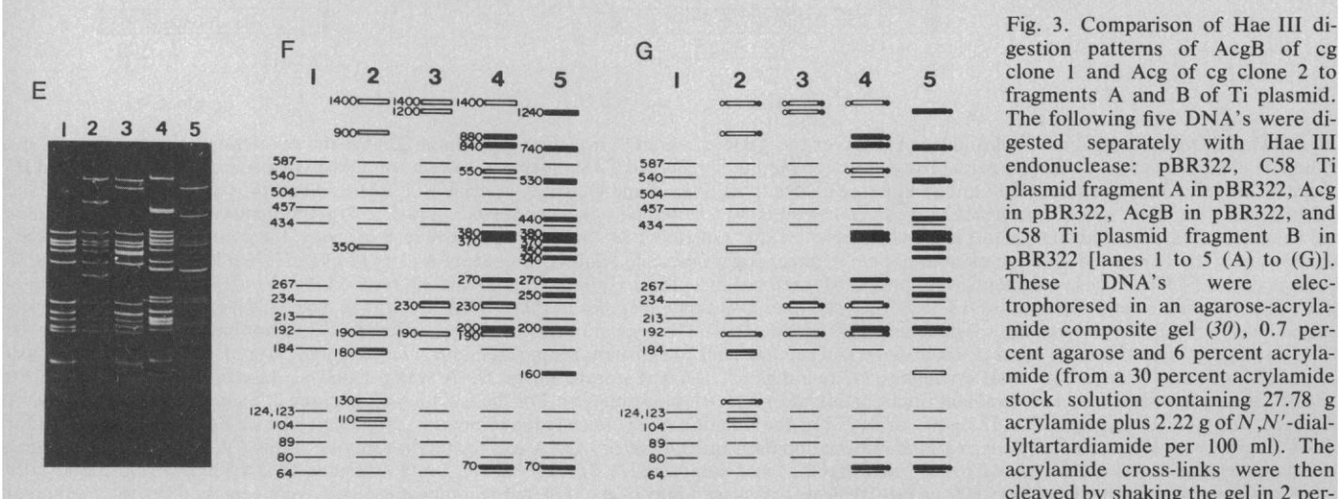




Fig. 3. Comparison of Hae III digestion patterns of AcgB of cg clone 1 and Acg of cg clone 2 to fragments A and B of Ti plasmid. The following five DNA's were digested separately with Hae III endonuclease: pBR322, C58 Ti plasmid fragment A in pBR322, Acg in pBR322, AcgB in pBR322, and C58 Ti plasmid fragment B in pBR322 [lanes 1 to 5 (A) to (G)]. These DNA's were electrophoresed in an agarose-acrylamide composite gel (*30*), 0.7 percent agarose and 6 percent acrylamide (from a 30 percent acrylamide stock solution containing 27.78 g acrylamide plus 2.22 g of *N*,*N*′-diallyltartardiamide per 100 ml). The acrylamide cross-links were then cleaved by shaking the gel in 2 percent periodic acid at 37°C for 15 minutes. The gel was denatured in 0.5*M* NaOH plus 1.5*M* NaCl for 10 minutes, neutralized in 0.5*M* tris (*p*H 7.5) plus 3*M* NaCl for 10 minutes, and transferred to nitrocellulose (Schleicher and Schuell) in 0.2*M* Na acetate (*p*H 4) at room temperature (*30*). (E) Ethidium bromide staining pattern of gel-separated Hae III fragments. The DNA blots were hybridized to radioactive probes specific for purified fragment A (A), purified fragment B (B), purified fragment Acg (C), or purified fragment AcgB (D). Results of hybridization patterns in (A) and (B) are summarized in (F). Fragment sizes are given in base pairs. White fragments hybridize to radioactive fragment A, and black fragments hybridize to radioactive fragment B. Results of hybridization patterns in (C) and (D) are summarized in (G). White and black fragments are the same as in (F). White circles represent fragments which hybridize to radioactive fragment Acg, and black circles, fragments which hybridize to radioactive fragment AcgB. Hybridization and wash conditions were as described in Fig. 2, except that the temperature was 37°C. (F) Hae III fragments of Acg or AcgB which are derived from fragment A or B of Ti plasmid. For example, in both Acg and AcgB there are fragments of 1400 and 190 bp which are identical in size to those derived from fragment A of Ti plasmid [compare (F), lanes 2, 3, and 4]. In AcgB there are fragments of 380, 370, 270, 200, and 70 bp which are identical to those in fragment B of Ti plasmid [compare (F), lanes 4 and 5]. The hybridization of the 70-bp fragment is not visible in the photograph shown here. The other fragments represent either true boundary fragments of T-DNA or boundaries in the cloning vehicle pBR322 (for example, Hae III fragments of 230 and 270 bp). By digestion with different restriction enzymes these two possibilities can be distinguished (data not shown). (G) Hae III fragments in Ti plasmid which contain sequences homologous to T-DNA boundary Hae III fragments of Acg and AcgB. Both Acg and AcgB hybridize to Hae III fragments of 1400, 900, 190, and 130 bp derived from Ti plasmid fragment A [(G), lane 2]. AcgB hybridizes to Hae III fragments of 1240, 740, 530, 380, 370, 270, 200, and 90 bp derived from Ti plasmid fragment B [(G), lane 5].

highly repeated plant DNA sequences and that the repeated sequences are rich in A-T since they are detected only under nonstringent hybridization conditions. Sequence analysis indicates that the first 150 bp from the Eco RI site of Acg are approximately 77 percent A-T. The plant DNA in Acg contains sequences unique to tobacco, since Acg does not hybridize to DNA from *Petunia*, *Arabidopsis*, *Agrobacterium*, or rat pancreas (data not shown).

### Sequence Analysis

The preliminary DNA sequence of the junction region of cg clone 1 (stippled region of AcgB as shown in Fig. 5A) has several interesting features (Fig. 6). In

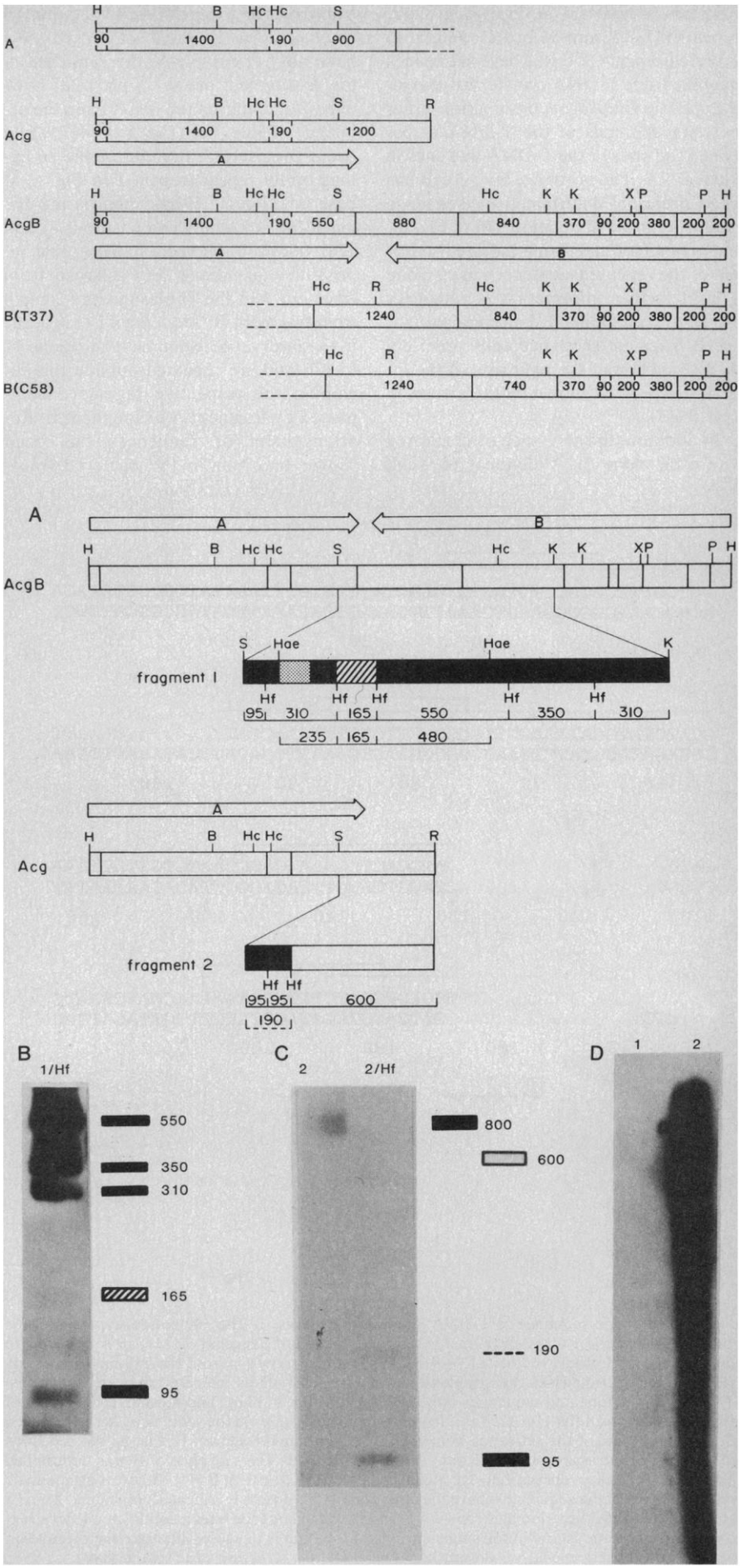



Fig. 4 (top). Hae III restriction maps of boundaries of T-DNA. The order of the Hae III restriction fragments analyzed in Fig. 3 was determined by comparing the sizes of the fragments produced after digestion with either Hae III or Hind III plus a second enzyme: Bam HI, Hinc II, Kpn I, Pst I, Eco RI, Sac II, or Xba I (diagrammed as *B*, *Hc*, *K*, *P*, *R*, *S*, or *X*, respectively). Boxes delineate Hae III sites except that the boundaries of clone AcgB are Hind III (*H*) or of Acg are Hind III and Eco RI. Arrows indicate which portions of Acg and AcgB hybridize to fragment A or B. The Hae III maps for fragment B from strains T37 and C58 are shown. Fig. 5 (bottom). Hinf I restriction maps of the boundaries of T-DNA and hybridization of T-DNA boundaries to total plant DNA. The Hae III map of fragments AcgB and Acg is reproduced from Fig. 4 except that the sizes of the fragments are not included. The Sac II/Kpn I fragment of AcgB (fragment 1) and the Sac II/Eco RI fragment of Acg (fragment 2) were further digested with Hinf I (*Hf*) and the order of the fragments was determined (*31*). (A) Hinf I restriction map of fragments 1 and 2. Fragment 1 was digested with Hinf I and electrophoresed on an agarose-acrylamide gel (see Fig. 3 legend for details) and hybridized to Ti fragments A and B (B) after blotting and transfer. Similarly, fragment 2 was digested with Hinf I and hybridized to Ti fragment A (C). The ethidium bromide staining patterns are diagrammed beside the autoradiogram of the hybridization pattern in (B) and (C). The diagram also indicates the extent of hybridization: black, strong hybridization; white, no hybridization; and crosshatched and stippled, weak hybridization. The band labeled 190 in (C) is a partial digestion product which was barely visible on ethidium bromide-stained gel but gave a positive hybridization signal. In contrast, the 165-bp fragment in (B) was visible but gave a weak hybridization signal. The extent of hybridization in (B) and (C) is also indicated on the Hinf I map in (A). In (D), fragments 1 and 2 were used as hybridization probes to total untransformed tobacco DNA digested with Eco RI. For (B) and (C) hybridization and washing were performed at 37°C. For (D) hybridization was carried out at 25°C and the filter was washed at 55°C in 3× SSC containing 0.1 percent SDS.

particular, there are several long direct repeats (13, 15, and 36 bp) as well as an inverted repeat (18 bp). These sequences may facilitate T-DNA transfer by providing specific sites for recombination either between the ends of the T-DNA or between the ends of the T-DNA and sites in plant DNA. The border of B in AcgB has been precisely determined by sequence analysis (arrow at position 176 in Fig. 6), but we cannot determine the precise origin of the repeated sequences in cg clone 1 until further information is obtained. However, preliminary data indicate that the B fragment sequence ends precisely at the junction of the right arm of the inverted repeat, as shown by the arrows in Fig. 6 (*18*).

In addition, the sequence of clone Acg (from cg clone 2) is identical to AcgB from position 1 to position 40 and diverges at position 41 (data not shown). We have not yet completed the sequence of the A fragment of the Ti plasmid; however, if it confirms the result from the sequence of Acg, then the A border would occur precisely at the end of the 36-bp-long direct repeat (repeat 1 in Fig. 6) at base position 40. What, then, is the origin of the DNA between positions 41 and 176? It would not seem unlikely that repeated sequences 1 and 3 arose from fragment A of the Ti plasmid and 2 and 4 from fragment B; sequences 1 to 4 occur in the interval between base positions 41 to 176 and are repeated outside this region in the respective regions encompassed by fragment A or fragment B. Rearrangement of sequences has been shown to occur in the bacterial transposon IS2 (*19*). The remainder of the DNA in the interval may also have been derived from regions A and B or even from plant DNA.

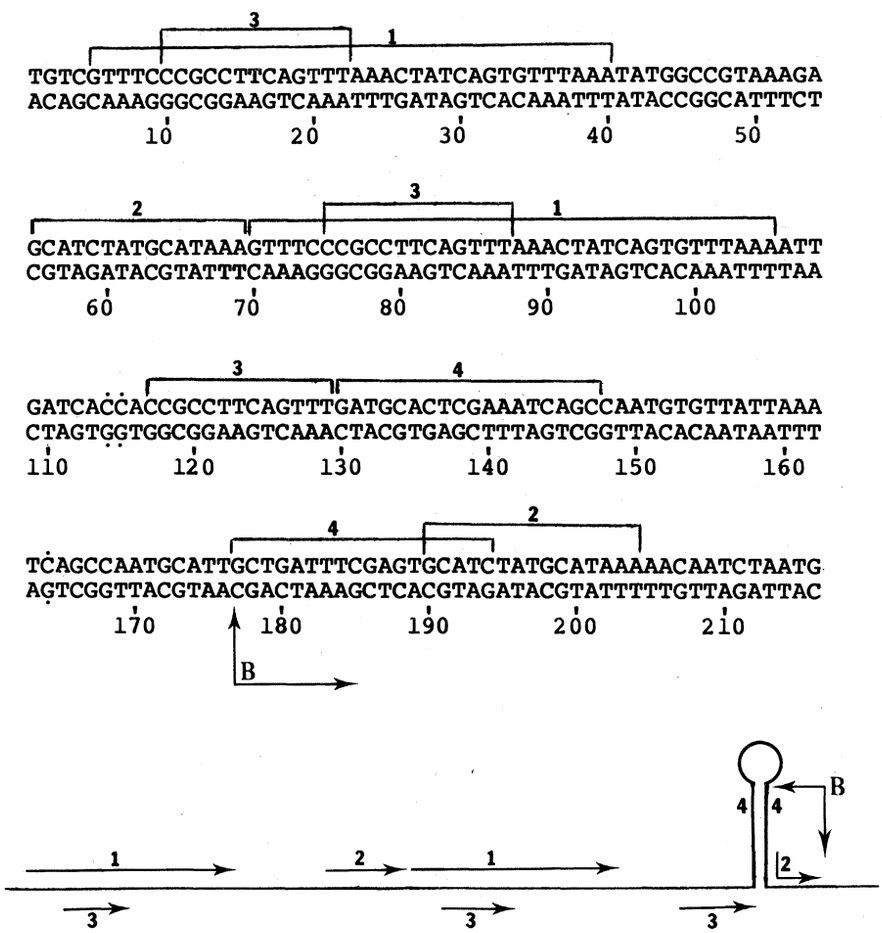



Fig. 6. Nucleotide sequence of T-DNA junction in cg clone 1. The nucleotide sequence of the junction of cg clone 1 was determined (*32*) by the method of Sanger *et al.* (*33*) after cloning into the single-strand phage vector M13 (*24*). The sequence shown is from the 310-bp Hinf I fragment, which encompasses the stippled area in Fig. 5. Brackets indicate repeated sequences. The sizes of the repeated sequences are: 1, 36 bp; 2, 15 bp; 3, 13 bp; and 4, an inverted repeat of 18 bp. Dots at positions 114, 115, and 164 indicate uncertainties in the sequence. Arrows labeled *B* indicate the end of the sequence homologous to Ti plasmid fragment B; that is, the sequence to the right of the arrows is identical to the sequence of B. The sequence of B was determined after labeling the appropriate Eco RI site in a clone of T37 fragment B (*34*). Below is a schematic representation of the sequence where the direction of the repeats is indicated; repeats 1, 2, and 3 are all in the same direction and repeat 4 is an inverted sequence which can form a loop structure. The schematic is drawn for single-stranded DNA and is meant to illustrate the orientation of the repeated sequences; we do not know whether this structure can form in vivo.

## Discussion

The boundaries of the T-DNA in the T37 teratoma were cloned in bacteriophage λ; one clone, cg clone 1, contains the left and right borders of the T-DNA linked together, and the other clone, cg clone 2, contains the right end of the T-DNA linked to repetitive plant DNA sequences. These data suggest that the T-DNA is organized as tandem repeats which are inserted into repetitive sequences of plant DNA. This model is supported by additional results obtained by Southern blotting analysis of the T-DNA of T37 teratoma (*14*); the 12.4-kbp Eco RI fragment (cg clone 1; see Fig. 2B) and the other internal Eco RI fragments of the T-DNA always give stronger hybridization signals than the 8.7-kbp Eco RI fragment (cg clone 2). The intensity of these bands suggests that there may be two to five tandem repeats of the T-DNA in this tumor cell line. It is not known whether the repetition of the T-DNA occurred before or after integration of the T-DNA. It will be interesting to compare other nopaline Ti plasmid–induced tumor cell lines to determine whether the T-DNA is always integrated in a tandem array. The analysis of cg clone 1 also shows that simple blot hybridization experiments are not sufficient to determine true boundaries of an integrated DNA molecule. It should be noted that although there is evidence that the T-DNA is found in the plant nucleus (*20*) and the data presented here also show that the T-DNA is linked to plant sequences, it has not been proved that the T-DNA is a stable part of a plant chromosome. Because of the low stringency required to show hybridization of Acg probe to plant DNA, it might be argued that formal homology or linkage to plant DNA has not been proved. However, the origin of the sequences from another source would be difficult to explain. It is possible that the T-DNA was once integrated but now exists as an independent replicon which retains a remnant of plant DNA sequences, or there may be two states of the T-DNA, one integrated and the other not. Experiments with undigested tumor DNA have failed to reveal the presence of single T-DNA replicons but have not excluded the possibility of polymeric circular T-DNA molecules (*14*).

The overall length of the T-DNA in in-

dependently isolated nopaline tumors is similar to that of the T37 teratoma analyzed in detail here and spans the region of the Ti plasmid from Hind III fragment B to fragment A (*14*). By contrast, in tumors induced by octopine Ti plasmids the T-DNA complement varies in different tumor cell lines (*13, 21*). In these cases the T-DNA always contains a particular region of the octopine Ti plasmid on the left side but may contain additional noncontiguous regions of Ti plasmid DNA on the right side (*13*). Thus, nopaline Ti plasmid–induced tumors may represent a simpler and more specific system for studying the mechanism of T-DNA formation.

The results presented here provide direct evidence that a specific piece of the Ti plasmid, the T region, can be transferred and integrated into plant cell DNA during tumor formation. Either the T region ends recombine before insertion into plant DNA, or the entire T region is excised and subsequently inserted to form the T-DNA (see Fig. 1). In either case, the "ends" of the T region are involved. It should be possible to understand the nature of the recombination event by comparing the nucleotide sequence of T-DNA junctions in cg clones 1 and 2 to the nucleotide sequence of the homologous regions of the Ti plasmid. The data obtained to date on the sequence of the junction of cg clone 1 reveal several direct and one inverted repeat. (It is possible that these structures are artifacts of cloning in *Escherichia coli*, but this seems unlikely.)

These repeated sequences may mediate DNA transfer by a mechanism analogous to that of transposable elements [reviewed in (*22*)]. The T-DNA exhibits one of the central properties of transposons by behaving as a discrete unit of DNA which can integrate into nonhomologous DNA sequences (*22*). It remains to be shown whether the T-DNA can also excise and mediate the formation of nearby deletions and inversions, as do transposable elements. The T-DNA represents a unique example of a transferable DNA element since it is able to cross biological barriers from a prokaryotic to a eukaryotic cell. This suggests there may be some mechanisms common to all DNA transfer. It is also possible that the T-DNA is involved in site-specific recombination with itself or with plant DNA in a manner analogous to that used in integration of bacteriophage into bacterial chromosomes. It will be interesting to compare the mechanism of T-DNA integration into plant DNA with the still to be determined mechanism for the movement of other genetic elements in plants, such as those first described in maize by McClintock (*23*).

### References and Notes


1. A. C. Braun and H. N. Wood, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 496 (1976).
2. J. Schell *et al.*, *Proc. R. Soc. London Ser. B* **204**, 251 (1979).
3. J. Schell and M. Van Montagu, in *Genomic Organization and Expression in Plants*, C. J. Leaver, Ed. (Plenum, New York, 1980), pp. 453–470; in *The Molecular Basis of the Infective Process*, H. Smith, J. J. Skehel, M. J. Turner, Eds. (Verlag Chemie, Berlin, in press); M. Van Montagu *et al.*, *Proc. R. Soc. London Ser. B*, in press.
4. N. Van Larebeke, G. Engler, M. Holsters, S. Van den Elsacker, I. Zaenen, R. A. Schilperoort, J. Schell, *Nature (London)* **252**, 169 (1974); N. Van Larebeke, C. Genetello, J. Schell, A. K. Hermans, J. P. Hernalsteens, M. Van Montagu, *ibid.* **255**, 742 (1975); N. Van Larebeke, C. Genetello, J. P. Hernalsteens, A. Depicker, I. Zaenen, E. Messens, M. Van Montagu, J. Schell, *Mol. Gen. Genet.* **152**, 119 (1977); B. Watson, T. C. Currier, M. P. Gordon, M. D. Chilton, E. W. Nester, *J. Bacteriol.* **123**, 255 (1975); M. D. Chilton, S. K. Farrand, R. Levin, E. W. Nester, *Genetics* **83**, 609 (1976).
5. C. Genetello, N. Van Larebeke, M. Holsters, M. Van Montagu, J. Schell, *Nature (London)* **265**, 561 (1977); A. Kerr, P. Manigault, J. Tempe, *ibid.* **264**, 560 (1977); M. Holsters, D. De Waele, A. Depicker, E. Messens, M. Van Montagu, J. Schell, *Mol. Gen. Genet.* **163**, 181 (1978); A. Petit, S. Delhaye, J. Tempe, G. Morel, *Physiol. Veg.* **8**, 205 (1970); G. Bomhoff, P. M. Klapwyk, H. C. Kester, R. A. Schilperoort, J. P. Hernalsteens, J. Schell, *Mol. Gen. Genet.* **145**, 171 (1976); A. Kerr and W. P. Roberts, *Physiol. Plant Pathol.* **9**, 202 (1976); A. L. Montoya, M. D. Chilton, M. P. Gordon, D. Sciaky, E. W. Nester, *J. Bacteriol.* **129**, 101 (1977).
6. A. Depicker, M. Van Montagu, J. Schell, *Nature (London)* **275**, 150 (1978).
7. D. Sciaky, A. L. Montoya, M. D. Chilton, *Plasmid* **1**, 238 (1978).
8. G. Engler, A. Depicker, R. Villarroel, R. Maenhout, M. Van Montagu, J. Schell, *J. Mol. Biol.*, in press.
9. A. Depicker, M. DeWilde, G. DeVos, R. DeVos, M. Van Montagu, J. Schell, *Plasmid* **3**, 193 (1980).
10. M. Holsters *et al.*, *ibid.*, p. 212.
11. M. D. Chilton, M. H. Drummond, D. J. Merlo, D. Sciaky, *Nature (London)* **275**, 142 (1978).
12. _____, A. L. Montoya, M. P. Gordon, E. W. Nester, *Cell* **11**, 263 (1977).
13. M. F. Tomashow, R. Nutter, A. L. Montoya, M. P. Gordon, E. W. Nester, *ibid.* **19**, 729 (1980).
14. M. Lemmers, M. DeBeuckeleer, A. Depicker, M. Holsters, P. Zambryski, J. P. Hernalsteens, M. Van Montagu, J. Schell, in preparation.
15. P. R. White and A. C. Braun, *Cancer Res.* **2**, 597 (1942).
16. A. C. Braun, *The Biology of Cancer* (Addison-Wesley, Reading, Mass, 1974).
17. F. R. Blattner *et al.*, *Science* **196**, 161 (1977).
18. This conclusion currently rests on a single sequencing analysis which indicates that the sequence of B contains an additional G residue inserted between positions 176 and 177 as compared to the sequence of AcgB (Fig. 6). A single G in the fragment B sequence at this position would allow the inverted repeat to be extended and the border with B would then not occur precisely at the junction of the inverted repeat.
19. D. Ghosal and H. Saedler, *Nature (London)* **275**, 611 (1978); *Cold Spring Harbor Symp. Quant. Biol.* **43**, 1193 (1978).
20. M.-D. Chilton, R. K. Saiki, N. Yadav, M. P. Gordon, F. Quetier, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 4060 (1980); L. Willmitzer, M. DeBeuckeleer, M. Lemmers, M. Van Montagu, J. Schell, in preparation.
21. M. DeBeuckeleer *et al.*, in preparation.
22. M. P. Calos and J. H. Miller, *Cell* **20**, 579 (1980).
23. B. McClintock, *Dev. Biol. Suppl.* **1**, 84 (1967).
24. B. Cordell, G. Bell, E. Tischer, F. DeNoto, A. Ullirch, R. Pictet, W. J. Rutter, H. M. Goodman, *Cell* **18**, 533 (1979).
25. T. Maniatis, R. C. Hardison, E. Lacy, J. Lauer, C. O'Connell, D. Quon, G. K. Sim, A. Efstratiadis, *ibid.* **15**, 687 (1978).
26. W. D. Benton and R. W. Davis, *Science* **196**, 180 (1977).
27. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
28. D. Denhardt, *Biochem. Biophys. Res. Commun.* **23**, 641 (1966).
29. H. M. Goodman and R. J. MacDonald, *Methods Enzymol.* **68**, 75 (1979).
30. J. Reiseer, J. Renart, G. R. Stark, *Biochem. Biophys. Res. Commun.* **85**, 1104 (1978).
31. H. O. Smith and M. L. Birnstiel, *Nucleic Acids Res.* **3**, 2387 (1976).
32. K. Kruger and A. Depicker, unpublished results.
33. F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977); F. Sanger and A. R. Coulson, *FEBS Lett.* **87**, 107 (1978).
34. A. M. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1980).


35. We are grateful to M. DeBeuckeleer for preparing plant DNA's. We thank S. Bock for computer editing of the manuscript and D. D. Moore and K. R. Yamamoto for critical readings. We thank all the members of our laboratories for helpful discussions. Supported by USDA grant 5901-0410-9-0374-0 and NIH grant CA 27424A, and in part by NIH grant CA14026. H.M.G. is an investigator of the Howard Hughes Medical Institute.

25 June 1980