Mol Gen Genet (1990) 224:248–256

MGG

Plant chromosome/marker gene fusion assay for study of normal and truncated T-DNA integration events

L. Herman*, A. Jacobs, M. Van Montagu, and A. Depicker

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

Received May 8, 1990

Summary. During *Agrobacterium tumefaciens* infection, the T-DNA flanked by 24 bp imperfect direct repeats is transferred and stably integrated into the plant chromosome at random positions. Here we measured the frequency with which a promoterless reporter gene is activated after insertion into the *Nicotiana tabacum* SR1 genome. When adjacent to the right or left T-DNA border sequences, at least 35% of the transformants express the marker gene, suggesting preferential T-DNA insertion (>70%) in transcriptionally active regions of the plant genome. When the promoterless neomycin phosphotransferase II (*nptII*) gene is located internally in the T-DNA, the activation frequency drops to 1% since gene activation requires T-DNA truncation. These truncation events in the *nptII* upstream region occur independently of the nature of the upstream sequence and of the T-DNA length. Deletion of the right border region prevents the detection of activated marker genes. Therefore, T-DNA truncation probably occurs after synthesis of a normal T-DNA intermediate during the transfer and/or integration process. In the absence of border regions, expression of the *nptII* selectable marker directed by the nopaline synthase promoter was detected in 1 out of 10^5 regenerated calli, suggesting the possibility that any DNA sequence from the Ti plasmid can be transformed into the plant genome, albeit at a low frequency.

Key words: *Agrobacterium tumefaciens* – T-DNA borders – Intergeneric gene transfer – pseudoborder – Insertional activation – Aberrant T-DNAs

Introduction

The transformation of plant cells by *Agrobacterium tumefaciens* is dependent on the insertion of a precisely delimited bacterial DNA sequence, the T-DNA, into the host genome (for a review see Gheysen et al. 1989). Upon contact with phenolic compounds released by wounded plant cells, the *Agrobacterium* transfer functions (*vir*) are induced (Stachel et al. 1985). The VirD1 and VirD2 proteins introduce single-stranded nicks in the lower strand of the 24 bp repeat sequence which flanks the transferred DNA region (Yanofsky et al. 1986; Wang et al. 1987). Subsequently, single-stranded linear molecules (T-strands) are generated by strand displacement with their 5′ and 3′ ends terminating at the nicked sites within the right and left T-DNA border repeats (Stachel et al. 1986; Albright et al. 1987). The right border repeat is absolutely required *in cis* for T-DNA transfer most probably because it directs the T-strand synthesis, whereas the left border repeat seems to act as a termination signal beyond which DNA sequences are normally not transferred (Wang et al. 1984).

The *overdrive* sequence at the right of the right border repeat is necessary to obtain high T-DNA transfer frequencies (Peralta et al. 1986). Consequently, the T-strand molecules become associated with several proteins in the so-called T-complexes. The VirE2 protein exhibits a single-stranded DNA-binding activity and probably plays a role in protecting the T-DNA strand during transfer from the bacterium to the plant (Christie et al. 1988; Citovsky et al. 1988; Das 1988; Dürrenberger et al. 1989). The VirD2 protein is covalently linked to the 5′ end of the T-strands and may play a role in the translocation of the T-complex to the plant cell nucleus by acting as a pilot protein (Young and Nester 1988; Herrera-Estrella et al. 1988, 1990). Thereafter, the T-DNA fragments are integrated at random positions in the plant genome (Ambros et al. 1986; Chyi et al. 1986; Wallroth et al. 1986).

DNA sequence analysis of T-DNA junctions cloned from genomes of transformed plants indicates that the junction points are located within or near the 24 bp imperfect repeats, which flank the transferred DNA region (Bakkeren et al. 1989). However, aberrant forms of T-DNA insertion are found in transformed plants with varying frequencies (1% to 40%), most commonly showing a deletion at one or both ends of the T-DNA (Spielmann and Simpson 1986; Deroles and Gardner 1988). It is not known how and when these truncated

* *Present address:* Rijkszuivelstation, Brusselsesteenweg 370, B-9230 Melle, Belgium

Offprint requests to: M. Van Montagu

T-DNAs are formed: during synthesis, transfer, or integration. It was suggested that truncated T-DNAs arise from the recognition of "pseudoborders" during T-strand synthesis in the bacterium (Hepburn and White 1985; Van Lijsebettens et al. 1986). This suggestion is based on correlations found between the truncation points as determined by genomic blotting and the positions of sequences resembling the 24 bp border repeat. The sequences of several aberrant T-DNA-plant junctions, however, do not reveal homology with the border consensus sequence (Gheysen et al. 1990). Therefore, even if pseudoborder sequences are at low frequency responsible for the generation of aberrant T-DNAs, other mechanisms resulting in random distribution of the endpoints must occur, such as: (i) the use of random nicks for the synthesis of shortened T-strands, (ii) digestion of the T-DNA ends by nucleases prior to integration, (iii) breakage of the T-DNA during the transfer process (i.e. in the bacterium or in the plant), and (iv) rearrangements generated during integration.

In this paper we analyse in more detail the process of T-DNA truncation. We have used the method of cocultivation of plant protoplasts with *Agrobacterium*, since this system allows detection and quantification of low frequency DNA transfer events. Firstly we determined the frequency of T-DNA integration into transcriptionally active regions of the plant genome using T-DNAs carrying a promoterless reporter gene next to the T-DNA border sequences. Secondly, we measured the frequency of T-DNA truncation by selecting for fusions of a marker gene located in the middle of the T-DNA with plant expression signals. Lastly, we determined the frequency of DNA transfer in the absence of right and left T-DNA border regions.

Materials and methods

Bacterial strains and plasmids. Plasmid constructions were transformed into *Escherichia coli* MC1061 (Casadaban and Cohen 1980). All recombinant DNA techniques were performed according to Maniatis et al. (1982).

Plasmid pGVL120 contains the bacterial spectinomycin (Sp) resistance gene and a promoterless *nptII* gene between octopine T-DNA borders (GenBank Accession No. M35007) (Fig. 2A) and the β-lactamase gene outside the T-DNA (Herman et al. 1986). pGVK50 is a cloning intermediate derived from pGVL120, with the Km^R gene of Tn*903* cloned in the *Sal*I site outside the T-DNA of pGVL120. Subsequently, pGVRBN was derived from pGVK50 by deleting the internal T-DNA *Hae*III-*Bgl*II segment and the *nptII* translation start codon and inserting the chimaeric hygromycin resistance gene (P*nos-hpt*-3′ *ocs*) in the *Cla*I site of the T-DNA (Fig. 1A and B). pGVRBH is derived from pGVL136 (Herman et al. 1986) by insertion of the coding sequence of the hygromycin resistance gene (*hpt*) with the 3′ end of the nopaline synthase (3′ *nos*) gene into the *Bam*HI site (Fig. 1A and B). pGVLBN is derived by insertion of the *Hin*dIII-*Apa*I fragment of pGVL120 containing the *nptII*-3′ *ocs* gene into the large *Hin*dIII-*Apa*I fragment of pGVL202

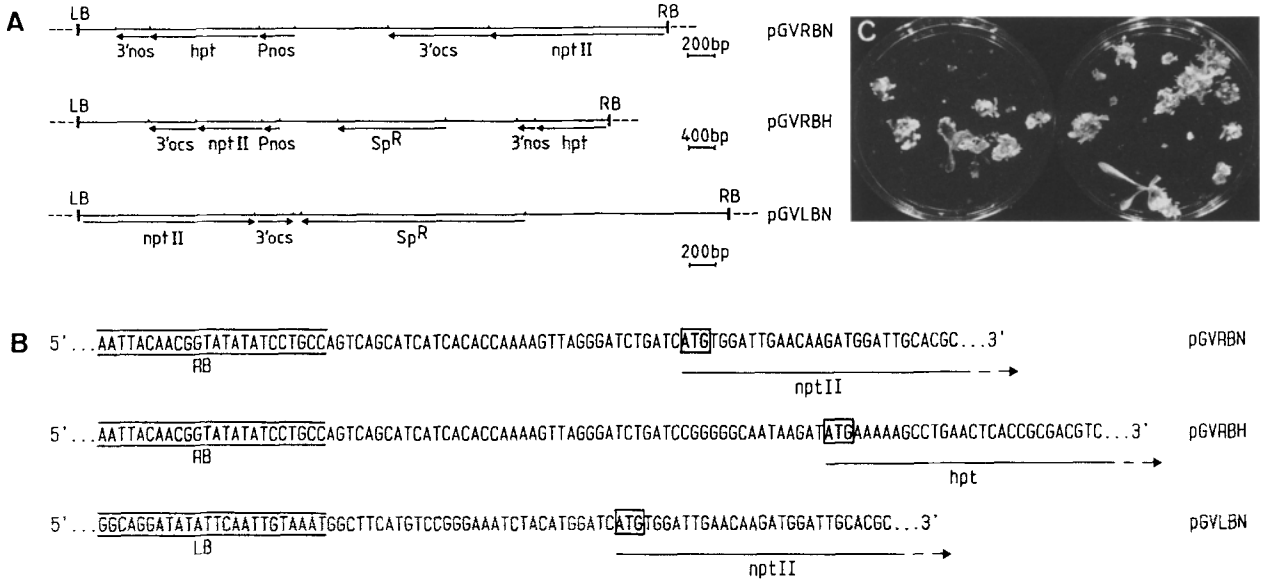

Fig. 1A–C. Insertional activation of a promoterless T-DNA gene adjacent to the right or left border. **A** Schematic representation of the T-region of pGVRBN, pGVRBH and pGVLBN. RB and LB indicate the position of the right and left border, respectively. The *dotted line* represents the flanking pBR322 DNA; Sp^R is a bacterial spectinomycin resistance marker; P*nos* indicates the nopaline synthase promoter; 3′ *nos* and 3′ *ocs* represent the 3′ ends of the nopaline synthase gene and the octopine synthase gene, respectively; *nptII* and *hpt* represent the coding sequences of the *nptII* and *hpt* genes, respectively. **B** Sequences of the *nptII* and *hpt* upstream regions adjacent to the border sequence. The *boxed nptII* translation start codon in pGVRBN is situated 37 bp from the right border and in pGVLBN 30 bp from the left border. The *boxed hpt* translation start codon in pGVRBH is located 52 bp from the right border. **C** Kanamycin selection of the transformed pGVRBN minicalli. Transformed minicalli with the pGVRBN T-DNA were selected on 25 μg/ml hygromycin and subsequently screened on 100 μg/ml kanamycin. The calli show different phenotypes; some are growing rapidly and form shoots, others are growing slowly; about half of them are not growing at all

(Herman et al. 1986); by insertion of the spectinomycin resistance gene as a *Bam*HI-*Bgl*II fragment from pGV300 (A. Caplan, personal communication) into the *Bam*HI site; and by insertion of a 192 bp *Sau*3A fragment containing the left T-DNA border sequence of pGV0201 (De Vos et al. 1981) upstream from the *nptII* sequence (Fig. 1A and B).

pGVL150 is a derivative of pGVL120 (Herman et al. 1986) and for convenience is renamed pGVL120I1 (Fig. 2A). pGVL120I2 (B. Claes, personal communication) is derived from pGVL150 by insertion of the chimaeric hygromycin resistance gene (P*nos-hpt*-3′*ocs*) in the *Cla*I site of the T-DNA (Fig. 2A). pGVL120Δ1 is a pGVK50 derivative with a deletion of the *Bcl*I fragment upstream from the *nptII* gene (Fig. 2A).

pGVL120ΔRB1 is a pGVK50 derivative with a substitution of the right border (*Eco*RI-*Bam*HI) by the small pBR322 *Eco*RI/*Bam*HI fragment (Fig. 3). pGVL120ΔRB2 is a pGVK50 derivative with a substitution of the *nptII* upstream region including the right border (*Eco*RI/*Bgl*II) by the small pBR322 *Eco*RI-*Bam*HI fragment (Fig. 3).

pGVL136ΔRB (Fig. 4A) is derived from pGVL136 (Herman et al. 1986) by substitution of the right border region (*Eco*RI-*Bam*HI) by the small *Eco*RI-*Bam*HI pBR322 fragment. pLGVneo2103 is described by Hain et al. (1985).

Conjugation and plant transformation. All plasmid constructions were mobilized to *Agrobacterium* C58C1Rif^R (pGV2260) by triparental mating and stabilized by cointegration through the homologous pBR322 segment into the resident pGV2260 virulence plasmid (Deblaere et al. 1985). Southern blot analysis of total *Agrobacterium* DNA was performed to confirm the structure and the integrity of the T-DNAs in the cointegrates (Dhaese et al. 1983).

Nicotiana tabacum var. SR1 was transformed by cocultivation of the different *Agrobacterium* strains with regenerating protoplasts (Depicker et al. 1985). Transformed *Agrobacterium* minicalli were selected on 100 μg/ml kanamycin or 25 μg/ml hygromycin.

Plant DNA preparations and Southern analysis. Total plant DNA was prepared from 20 g callus tissue as described by Lemmers et al. (1980). Digested DNA was separated in 0.8% agarose gels. Transfer to nylon membranes and hybridizations were carried out according to the recommendations of the manufacturer (Amersham, UK). Probes were prepared with a Random Primed Labeling Kit (Boehringer, Mannheim, FRG).

Results

The frequency of insertional activation of a promoterless T-DNA gene, located next to the T-DNA border, suggests preferential insertion in transcriptionally active regions of the plant genome

André et al. (1986) and Teeri et al. (1986) demonstrated that the activation of a promoterless *nptII*-coding sequence adjacent to the T-DNA right border sequence is possible after plant cell transformation. We were interested in determining the frequency of this phenomenon. This would allow us not only to gain insight into the T-DNA gene-tagging efficiency but also to interpret the activation frequency of an internally located promoterless marker gene (Herman et al. 1986; Gheysen et al. 1990).

Therefore, T-DNAs were constructed with a promoterless neomycin phosphotransferase II (*nptII*) or hygromycin phosphotransferase (*hpt*) gene adjacent to the T-DNA border sequences. Plasmids pGVRBN and pGVRBH were constructed as described in the Materials and methods and are shown in Fig. 1A. The pGVRBN T-DNA contains the P*nos-hpt* gene as a selectable marker and the *nptII*-coding sequence with its start codon located 37 bp from the right border repeat (Fig. 1A). Similarly, pGVRBH contains the P*nos-nptII* gene as a selectable marker and the coding sequence for the *hpt* gene, starting 52 bp from the right border repeat (Fig. 1A). These T-DNA constructions were recombined in the *A. tumefaciens* plasmid pGV2260 (Deblaere et al. 1985) and used in co-cultivation experiments with protoplasts of *N. tabacum* SR1.

The minicalli co-cultivated with pGVRBN were selected for the expression of the hygromycin resistance gene. Subsequently, 600 hygromycin-resistant calli were screened on plates containing 100 μg/ml kanamycin. As shown in Fig. 1C, these calli showed different growth rates, corresponding to different levels of NPTII activity; 35% of the calli grew normally, 10%–15% hardly grew but did not die after two subsequent transfers, and 50%–55% were clearly kanamycin-sensitive (Table 1A). Kanamycin selection of regenerating tobacco protoplasts has proven to be a very reliable low-background method to identify plant clones expressing the *nptII* gene (Herman et al. 1986; André et al. 1986; Gheysen et al. 1990). The transformants contain variable amounts of NPTII activity with proteins of normal or enlarged molecular weight indicating transcriptional or translational fusions to plant-encoded expression signals (Koncz et al. 1989). Therefore, we conclude that at least 50% of the transformants contain a T-DNA with its *nptII*-coding sequence at the right border fused to plant expression signals active in callus.

The pGVRBH-transformed minicalli were selected for expression of the *nptII* selectable marker. Subsequently, 300 calli were screened for expression of the *hpt* resistance gene on plates containing 25 μg/ml hygromycin; 49% of these were found to be hygromycin-resistant (Table 1A). Thus, the *hpt* promoterless marker gene at the right T-DNA border is activated in 49% of the transformed calli, confirming the result obtained with the *nptII* marker gene.

We also determined the frequency of insertional activation of a promoterless gene placed close to the left border. Since the transcribed region in the plant genome is asymmetric with a complexed coding strand and an uncomplexed non-coding strand, it could be hypothesized that the polar VirD2-directed intermediate T-structure would preferentially react with the non-coding

Table 1. Selection of kanamycin-resistant plant clones after transformation with *Agrobacterium*, containing the *nptII* gene in different configurations

Agrobacterium carrying the plasmid construct	Colonies screened		Number of agarose discs screened[a]	Marker gene insertional activation frequency (%) per colony	
	Transformed	Regenerated		Transformed	Regenerated
A. pGVRBN	600			50	
pGVRBH	300			49	
pGVLBN		300			26
B. pGVL120			20		0.2
pGVL120I1			20		0.2
pGVL120I2			20		0.2
pVGL120Δ1			20		0.2
pGVL120ΔRB1			20		<0.001
pGVL120ΔRB2			20		<0.001
C. pGVL136		200			50
pGVL136ΔRB			20		0.1
pGVneo2103(ΔRB, ΔLB)			20		0.01

[a] Four days after co-cultivation, plants cells were embedded in agarose discs and selection was applied by the addition of 50 μg kanamycin (Km)/ml in the liquid medium (Depicker et al. 1985). Five weeks later, the agarose discs were placed on solidified B5 agar medium containing 50 μg Km/ml. The plates were scored for the number of colonies ten weeks after protoplast isolation. When grown in non-selective conditions, approximately 10^4 minicalli regenerated per agarose disc

strand of the target DNA. This would mean that insertional activation would be more frequent when a promoterless marker gene is located at the right rather than the left border.

Plasmid pGVLBN was constructed as described in the Materials and methods and is shown in Fig. 1A. The pGVLBN T-DNA contains the *nptII*-coding sequence starting 30 bp from the left border repeat. Since there is no additional selectable marker present, the protoplasts co-cultivated with pGVLBN were regenerated without selection. Afterwards, 300 calli were screened for the expression of the kanamycin resistance gene; 26% were found to be kanamycin resistant (Table 1A). Assuming a high transformation frequency (50%) which is commonly found in our co-cultivation experiments, we estimated that the frequency of insertional activation of the *nptII*-coding sequence next to the left border occurs in half of the transformed calli. This co-cultivation and screening experiment was repeated three times and each time the same activation frequency was obtained. The high frequency of insertional activation of a promoterless gene at both the left and right sides of the T-DNA suggests that there is a preferential insertion of the T-DNA into actively transcribed regions of the plant genome, and also that the insertion event does not show polarity.

The activation frequency of a promoterless nptII *gene, located internally in the T-DNA, stays constant for T-DNAs of different lengths and different* nptII *upstream sequences*

Selecting for a kanamycin-resistant phenotype after transformation of plant cells with a promoterless *nptII* gene located in the middle of the T-DNA can be used as a selection method for shortened T-DNA insertions. The *nptII*-coding sequence is activated by truncation of the T-DNA and subsequent fusion of the resistance gene with expression signals of the plant genome (Gheysen et al. 1990). To distinguish different possible mechanisms for the generation of truncated T-DNAs, we compared the insertional activation frequency of the *nptII*-coding sequence located in T-DNAs of different lengths and/or with a different *nptII* upstream DNA sequence.

Therefore, several pGVL120 derivatives were constructed as shown in Fig. 2A. The pGVL120 T-DNA is 6425 bp long and the distance between the start codon of the *nptII* gene and the right border sequence is 3670 bp (Herman et al. 1986). The pGVL120Δ1, pGVL120I1, and pGVL120I2 T-DNAs are 4775 bp, 7858 bp, and 9463 bp in length, respectively, with the *nptII*-coding sequence separated from the right border by 2010 bp in the first and 5093 bp in the two latter constructions.

After transformation of *N. tabacum* SR1 protoplasts with the pGVL120, pGVL120Δ1, pGVL120I1, and pGVL120I2 T-DNA constructions, the minicalli were grown with or without selection, as is shown in Fig. 2B for the pGVL120I2 transformation. Without selection, approximately 10^4 minicalli regenerated per agarose disc, whereas after hygromycin selection approximately 50% of the calli grew, indicating a high transformation frequency. After kanamycin selection, approximately 20 calli were found per agarose disc in all the different transformations; this means that the *nptII* marker is activated in 0.2% of the regenerated minicalli (Table 1B). This result was confirmed in four independent experiments.

The probability that the level of expression of the promoterless marker gene is sufficient to survive the se-

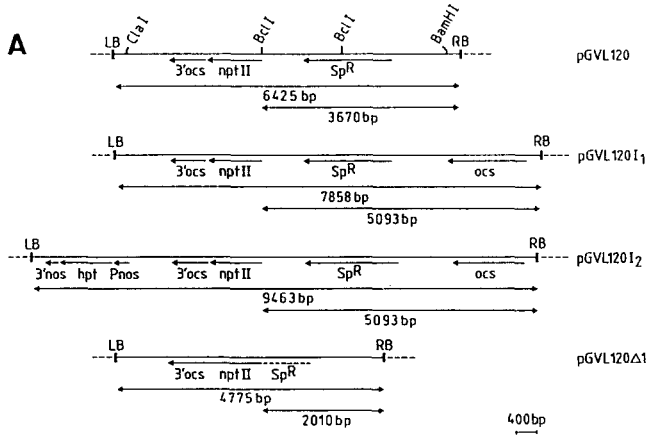

Fig. 2A and B. Transformation with T-DNAs containing an internal, promoterless *nptII* gene. **A** Schematic representation of the T-region derivatives of pGVL120. The length of the total T-DNA and the distance between the *nptII*-coding sequence and the right border are indicated below each construct. Only the restriction sites relevant for the cloning are indicated (see the Materials and methods). *ocs* indicates the octopine synthase gene with its own promoter and 3′ end signals. **B** Minicalli transformed with the pGVL120I2 T-DNA and grown with or without selective pressure, as indicated below each plate. *Nicotiana tabacum* SR1 protoplasts were co-cultivated with *Agrobacterium* and embedded to a density of approximately 10^4 minicalli per agarose disc

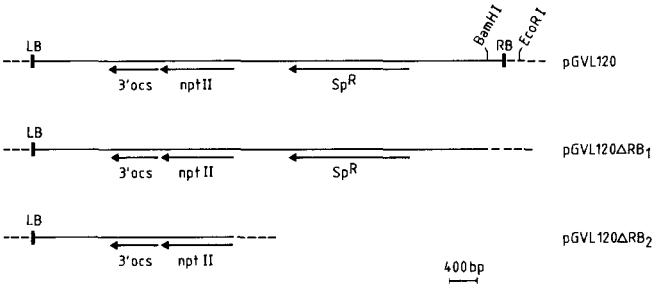

Fig. 3. Schematic representation of pGVL120 and derived constructions with deleted right border region. In pGVL120ΔRB1 and pGVL120ΔRB2 only the fragments corresponding to the T-DNA of pGVL120 are presented. Only the restriction sites relevant for the cloning are indicated (see the Materials and methods)

lection is assumed to decrease with the distance from the border sequence, since increasing the spacing enhances, for instance, the number of intervening translational start and stop codons. The T-DNA has to be truncated to delete these interfering translational signals and thus allow the expression of the internally located T-DNA marker. Since the insertional activation frequency is two orders of magnitude lower for the internally located promoterless *nptII* marker than for the *nptII* gene adjacent to the border (0.2% versus 25% of the regenerated calli), we estimated that truncation of the *nptII* upstream region occurs in approximately 1% of the transferred T-DNAs.

To explain this phenomenon of truncation the use of pseudoborders has been postulated (Hepburn and White 1985; Van Lijsebettens et al. 1986). However, the involvement of pseudoborders is rather unlikely since different *nptII* upstream sequences have no pronounced effect on the observed truncation frequency in a particular T-DNA segment. Also, five truncated T-DNA/plant junctions were cloned, sequenced, and compared with the nucleotide sequence of the T-DNA, but none showed a truncation point close to a pseudoborder (Gheysen et al. 1990). A mechanism, involving exonuclease digestion from the ends can be excluded since the truncation frequency would be inversely correlated with length. Our results favour the hypothesis that shortened T-DNA insertions are generated by breakage at some stage during the transfer and/or integration process.

The right border region is necessary for the efficient transfer of truncated T-DNAs

The presence of the right border region is a prerequisite for efficient T-DNA transfer to the plant genome. However, it is not known whether the right border region is also instrumental in the formation of aberrant T-DNAs. Study of the intermediate T-strand structures showed that all possible T-strands can be found between four nicked sites in the four border repeats (Stachel et al. 1987), suggesting that fortuitous nicks could be used for the synthesis of shortened T-DNAs. If indeed truncated T-DNAs are synthesized from randomly occurring nicks, then the frequency of occurrence will be the same whether or not the right border sequence is present. If, on the other hand, the truncated T-DNAs are derived from the normal T-DNA intermediates, then the frequency of their occurrence will decrease when the right border is deleted.

As shown in Fig. 3, a deletion of 280 bp removes only the right border and the overdrive sequence in pGVL120ΔRB1, whereas all T-DNA sequences, starting 9 bp upstream from the *nptII* translation codon and including the right border and the overdrive are deleted in pGVL120ΔRB2. Co-cultivation of protoplasts with the deletion derivatives, pGVL120ΔRB1 or pGVL120ΔRB2, showed that no kanamycin-resistant callus could be obtained (Table 1B), whereas transformation with the pGVL120 T-DNA yields kanamycin-resistant calli with a frequency of 0.2%. This means that the presence of the right border region is required for the detection of truncated T-DNAs and, therefore, favours the hypothesis that truncation occurs after synthesis of a normal T-DNA intermediate during the transfer and/or integration process.

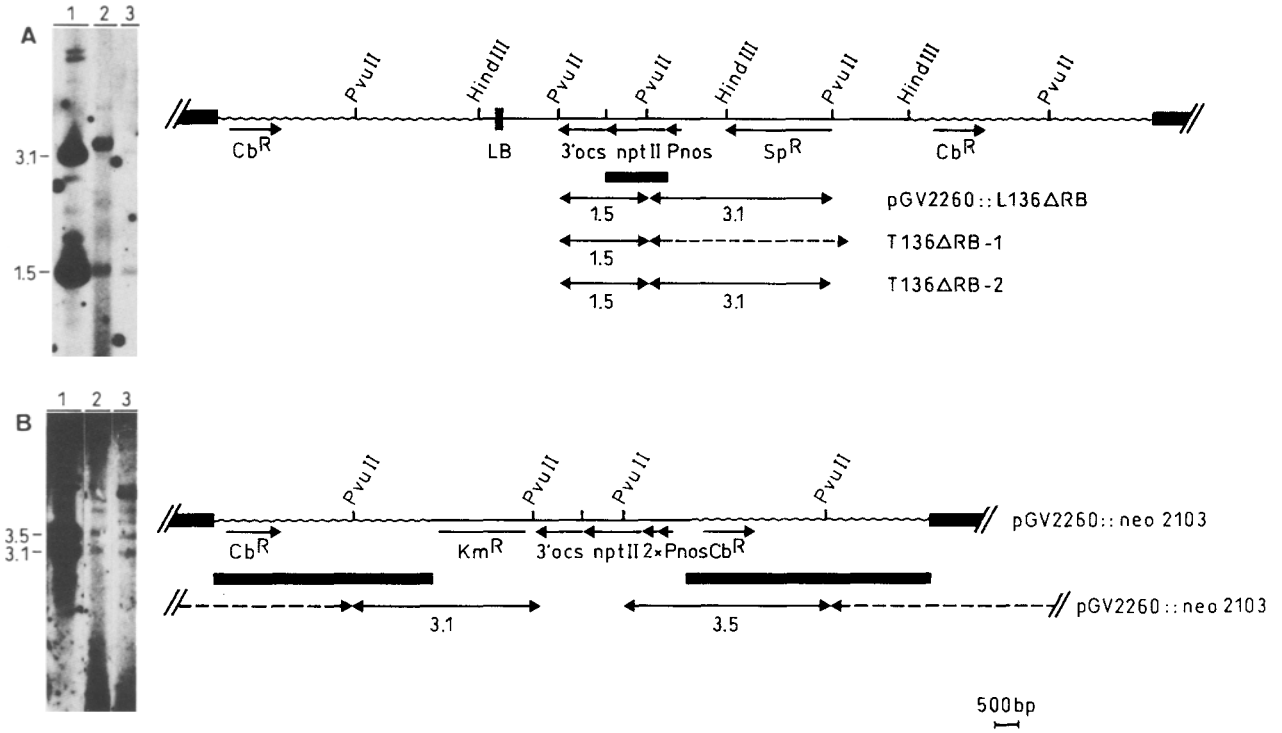

Fig. 4A and B. Genomic characterization of kanamycin-resistant plant clones, transformed with the P*nos-nptII* gene derived from the cointegrate plasmids **A** pGV2260::L136ΔRB and **B** pGV2260::neo2103. Lanes 1 in panels **A** and **B** show a control *Pvu*II digest of 100 ng of the infecting *Agrobacterium* DNA containing the pGV2260::136ΔRB cointegrate and the pGV2260::neo2103 cointegrate, respectively; lanes 2 and 3 contain 10 μg of DNA digested with *Pvu*II from plant clones transformed with pGV2260:136ΔRB (T136ΔRB-1 and T136ΔRB-2) and pGV2260::neo2103 (Tneo2103-1 and Tneo2103-2), respectively. The filter in **A** was hybridized with the pKC7 *Bgl*II-*Sma*I fragment which carries the *nptII*-coding sequence; the filter in **B** was hybridized with a pBR322 probe. The size of the hybridizing fragments is indicated in kb. The scheme at the right represents the cointegrate structure and the sizes of the hybridizing fragments in the experiment shown. The homology with the probes is represented by a *black bar*. The *wavy lines* and the *bold lines* represent pBR322- and Ti plasmid-homologous sequences, respectively

DNA transfer from the Agrobacterium *Ti plasmid to the plant genome can occur in the absence of the right and/or left border sequences*

Deletion of the left border repeat does not detectably affect T-region transfer, whereas deletion of the right border has been claimed either to attenuate or to abolish T-DNA transfer to the plant genome (Joos et al. 1983; Wang et al. 1984; Jen and Chilton 1986). Without the right border region no insertional activation of a promoterless marker gene could be detected and, therefore, it was important to determine more precisely the frequency of DNA transfer in the absence of border regions.

Different plasmids were constructed with a chimeric *nptII* gene under the control of the nopaline synthase promoter (P*nos*). pGVL136 contains the P*nos-nptII* gene flanked by both border sequences (Herman et al. 1986); pGVL136ΔRB is derived from this construct by deleting the right border region (Fig. 4A); and in pLGVneo2103 (Hain et al. 1985) no border regions are present (Fig. 4B).

After transformation with the pGVL136ΔRB construction, kanamycin-resistant calli were obtained with a frequency of 0.1% of the regenerated calli (Table 1C). This value is 500 times lower than the transformation frequency obtained when both borders are present, indicating that DNA transfer still occurs in the absence of the right border region but inefficiently. Two kanamycin-resistant plant lines, T136ΔRB-1 and T136ΔRB-2, were analysed by Southern hybridization. As shown in Fig. 4A, the internal 1.5 kb *Pvu*II fragment overlapping the *nptII* gene is present in both cell lines, whereas the 3.1 kb *Pvu*II fragment upstream from the *nptII* gene is only present in the T136ΔRB-2 plant line and not in the T136ΔRB-1 plant line. Since the 1.2 kb *Pvu*II-*Hin*dIII fragment upstream from the *nptII* gene is also lacking in the T136ΔRB-1 line (data not shown), the right endpoint of the insertion is located within this fragment close to the *nptII* gene. The right endpoint of the insertion in the T136ΔRB-2 line is located more than 2 kb upstream from the *nptII* gene within the pBR322 sequences or within Ti plasmid sequences. By hybridizing the genomic digests of both plant lines with a pBR322 probe, the left endpoints could be localized. In T136ΔRB-1, pBR322 sequences derived from the left repeat give a hybridization signal, but the left pBR322 *Pvu*II site is not present. In T136ΔRB-2, the *Hin*dIII fragment spanning the *nptII* gene is present as internal fragment but also in this line the pBR322 *Pvu*II site is no longer present. Thus, both transferred DNA fragments have their left endpoint in the pBR322 sequences upstream from the left border.

If both borders are omitted, as in pLGVneo2103,

the transfer frequency of the P*nos-nptII* gene drops to 0.01% (Table 1C). This indicates that the left border in the absence of the right border enhances the DNA transfer of adjacent sequences from *Agrobacterium* to the plant cell by at least one order of magnitude. Possibly the nicked border sequence stimulates the formation of intermediate T-DNA structures or allows processing of randomly initiated T-strands. Southern hybridizations of two kanamycin-resistant plant lines transformed with pLGVneo2103 indicate that indeed the P*nos-nptII* gene from the infecting strain is integrated with a variable amount of flanking pBR322 DNA. Figure 4B shows the result of the hybridization of a pBR322 probe with a *Pvu*II digest of both plant DNAs. Clearly, the 3.1 and 3.5 kb fragments are present whereas the flanking DNA differs in the two plant lines. This result seems to imply that any DNA sequence from the Ti plasmid can be transferred to the plant genome albeit with a low frequency.

Discussion

The frequency of T-DNA insertions within transcriptional units in the plant genome was determined by measuring the proportion of transformants in which a promoterless marker gene adjacent to the right border is activated. When the *nptII*-coding sequence is located 37 bp from the right border, 35% to 50% of the transformed calli express the *nptII* gene. A similar frequency for insertional gene activation was found for the *hpt*-coding sequence located 52 bp from the right border sequence. Also when the *nptII*-coding sequence is located next to the left border, close to 50% of the calli contain the activated marker gene. Therefore, as many as 70% to 100% of the transformed calli contain at least one of their T-DNA insertions in a transcriptionally active region with a transcript running into the T-DNA from either the left or the right. The frequency of T-DNA insertion in transcribed plant sequences is remarkably high, especially since only chromosome/marker gene fusions in the correct transcriptional orientation or in the appropriate translational reading frame will be detected. However, it should be noted that transformed calli often contain more than one T-DNA and, therefore, the chance that a particular T-DNA will integrate into a transcribed region is lower than the observed marker activation frequency. Preferential insertion of the T-DNA in transcriptionally active regions, as suggested by our results, has also been observed by Koncz et al. (1989). Therefore, T-DNA integration is random in respect of sequence specificity (Ambros et al. 1986; Chyi et al. 1986; Wallroth et al. 1986) but shows a clear bias for open transcribed chromatin. The open chromatin structure might facilitate recombination and serve as a preferential target for integration of the T-intermediate. Similarly, retroviral integration sites in transgenic Mov mice frequently map in transcribed DNA regions near DNase I hypersensitive sites (Mooslehner et al. 1990).

When the promoterless *nptII* gene is located internally in the T-DNA, kanamycin-resistant calli can arise only by breakage and subsequent fusion of the *nptII* gene with transcriptionally active plant DNA sequences. The frequency with which kanamycin-resistant calli can be selected is therefore the product of the transformation frequency, the truncation frequency, and the frequency of gene activation. Activated T-DNA insertions are detected in our selection system in 0.2% of the regenerated minicalli. Since we select only those aberrant T-DNA insertions which cause active *nptII* fusions, the truncation frequency anywhere in the T-DNA is much higher and, over the entire T-DNA, probably comparable to reported frequencies: 12% of the transformants after co-cultivation (Van Lijsebettens et al. 1986), at least 25% of the root cultures after transformation with *Agrobacterium rhizogenes* (Simpson et al. 1986), and between 30% and 40% of the transferred T-DNA copies after leaf disc infection (Spielmann and Simpson 1986; Deroles and Gardner 1988).

Based on our results, we favour the hypothesis that shortened T-DNA insertions are generated by breakage at some stage during the transfer or integration process, most probably after synthesis of a normal T-DNA intermediate. There are two lines of evidence supporting this hypothesis; firstly, no truncated T-DNAs could be detected when the right border region of pGVL120 was deleted, suggesting that truncation occurs after synthesis of a normal intermediate. Secondly, truncation occurs with equal frequency in different DNA sequences and at different distances from the right border, indicating that truncation is a random process and not sequence-dependent. The equal distribution of the truncations over the right and left border regions of the T-DNA (Deroles and Gardner 1988) is in agreement with this mechanism for generation of truncated T-DNAs.

We have tried to assess the relative importance of the right and left border for the frequency of transfer of DNA to the plant genome. As compared to the normal transformation frequency, deletion of the right border lowers the transfer frequency 500-fold, which means that 1 out of 1000 regenerated calli is transformed with the selective marker gene. This result confirms and extends the observations of previous studies which also detected reduced transfer of selected sequences when the right T-DNA border was deleted (Joos et al. 1983; Wang et al. 1984; Caplan et al. 1985; Hepburn and White 1985; Jen and Chilton 1986).

Here, we report for the first time that DNA transfer from the Ti plasmid of *Agrobacterium* to the plant genome can also occur in the absence of both the left and the right border regions. In about 1 out of 10^5 co-cultivated calli, insertion of the *nptII* selection marker is detected. Genomic characterization of two kanamycin-resistant transformed plant lines confirmed the insertion of the *nptII* sequence together with variable amount of adjacent sequences. These results suggest a low-frequency transfer of arbitrary *Agrobacterium* DNA to the plant genome, possibly due to the uptake of naked DNA, but more likely to a conjugative process. In this respect, Buchanan-Wollaston et al. (1987) reported that in a genetic background where *vir* genes are expressed the mobilization functions of the wide host-range plas-

mid pRSF1010 can mediate the efficient transfer of the *oriT*-adjacent DNA sequences from *Agrobacterium* into plant cells. More recently, Heinemann and Sprague (1989) have shown that several conjugative systems consisting of *tra* functions, the Mob protein, and *oriT* can mobilize *oriT*-linked DNA sequences from *E. coli* to yeast. Thus, we propose that the Ti plasmid *vir*-encoded functions also transmit single-stranded DNA molecules originating from the Ti plasmid *oriT* locus or maybe even from random nicks in *vir*-induced agrobacteria. This occurs only at a low frequency possibly because these single-stranded DNAs are inefficiently complexed or because the derived intermediates lack signals for transport and/or integration in the plant genome.

Acknowledgements. We acknowledge the contributions of A. Caplan, M. Van Lijsebettens, A. Herrera-Estrella, G. Gheysen, G. Angenon, and I. Ingelbrecht for critical reading and suggestions. We are also grateful to A. Van Kerrebroeck for help with the cloning experiments, M. De Cock for preparing the manuscript, and K. Spruyt, S. Van Gijsegem, and V. Vermaercke for preparation of figures and drawings. This work was supported by grants from the Services of the Prime Minister (U.I.A.P. 120C0187). LH is indebted to the I.W.O.N.L. for a post-doctoral fellowship.

References

Albright LM, Yanofsky MF, Leroux B, Ma S, Nester EW (1987) Processing of the T-DNA of *Agrobacterium tumefaciens* generates border nicks and linear, single-stranded T-DNA. J Bacteriol 169:1046–1055

Ambros PF, Matzke AJM, Matzke MA (1986) Localization of *Agrobacterium rhizogenes* T-DNA in plant chromosomes by in situ hybridization. EMBO J 5:2073–2077

André D, Colau D, Schell J, Van Montagu M, Hernalsteens J-P (1986) Gene tagging in plants by a T-DNA insertion mutagen that generates APH(3′) II-plant gene fusions. Mol Gen Genet 204:512–518

Bakkeren G, Koukolíková-Nicola Z, Grimsley N, Hohn B (1989) Recovery of *Agrobacterium tumefaciens* T-DNA molecules from whole plants early after transfer. Cell 57:847–857

Buchanan-Wollaston V, Passiatore JE, Cannon F (1987) The *mob* and *oriT* mobilization functions of a bacterial plasmid promote its transfer to plants. Nature 328:172–175

Caplan AB, Van Montagu M, Schell J (1985) Genetic analysis of integration mediated by single T-DNA borders. J Bacteriol 161:655–664

Casadaban MJ, Cohen SN (1980) Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J Mol Biol 138:179–207

Christie PJ, Ward JE, Winans SC, Nester EW (1988) The *Agrobacterium tumefaciens virE2* gene product is a single-stranded-DNA-binding protein that associates with T-DNA. J Bacteriol 170:2659–2667

Chyi YS, Jorgensen RA, Goldstein D, Tanksley SD, Loaiza-Figueroa F (1986) Locations and stability of *Agrobacterium*-mediated T-DNA insertions in the *Lycopersicon* genome. Mol Gen Genet 204:64–69

Citovsky V, De Vos G, Zambryski P (1988) Single-stranded DNA binding protein encoded by the *virE* locus of *Agrobacterium tumefaciens*. Science 240:501–504

Das A (1988) *Agrobacterium tumefaciens virE* operon encodes a single-stranded DNA-binding protein. Proc Natl Acad Sci USA 85:2909–2913

De Vos G, De Beuckeleer M, Van Montagu M, Schell J (1981) Restriction endonuclease mapping of the octopine tumor inducing pTiAch5 of *Agrobacterium tumefaciens*. Plasmid 6:249–253

Deblaere R, Bytebier B, De Greve H, Deboeck F, Schell J, Van Montagu M, Leemans J (1985) Efficient octopine Ti plasmid-derived vectors for *Agrobacterium*-mediated gene transfer to plants. Nucleic Acids Res 13:4777–4788

Depicker A, Herman L, Jacobs A, Schell J, Van Montagu M (1985) Frequencies of simultaneous transformation with different T-DNAs and their relevance to the *Agrobacterium*/plant cell interaction. Mol Gen Genet 201:477–484

Deroles SC, Gardner RC (1988) Analysis of the T-DNA structure in a large number of transgenic petunias generated by *Agrobacterium*-mediated transformation. Plant Mol Biol 11:365–377

Dhaese P, De Greve H, Gielen J, Seurinck J, Van Montagu M, Schell J (1983) Identification of sequences involved in the polyadenylation of higher plant nuclear transcripts using *Agrobacterium* T-DNA genes as models. EMBO J 2:419–426

Dürrenberger F, Crameri A, Hohn B, Koukolíková-Nicola Z (1989) Covalently bound VirD2 protein of *Arabidopsis thaliana* protects the T-DNA from exonucleolytic degradation. Proc Natl Acad Sci USA 86:9154–9158

Gheysen G, Herman L, Breyne P, Van Montagu M, Depicker A (1989) *Agrobacterium tumefaciens* as a tool for the genetic transformation of plants. In: Butler LO, Harwood C, Moseley BEB (eds) Genetic transformation and expression. Intercept, Andover, pp 151–174

Gheysen G, Herman L, Breyne P, Gielen J, Van Montagu M, Depicker A (1990) Cloning and sequence analysis of truncated T-DNA inserts from *Nicotiana tabacum*. Gene, in press

Hain R, Stabel P, Czernilofsky AP, Steinbiss H-H, Herrera-Estrella L, Schell J (1985) Uptake, integration, expression and genetic transmission of a selectable chimaeric gene to plant protoplasts. Mol Gen Genet 199:161–168

Heinemann JA, Sprague GF Jr (1989) Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. Nature 340:205–209

Hepburn AG, White J (1985) The effect of the right terminal repeat deletion on the oncogenicity of the T-region of pTiT37. Plant Mol Biol 5:3–11

Herman LMF, Van Montagu M, Depicker AG (1986) Isolation of tobacco DNA segments with plant promoter activity. Mol Cell Biol 6:4486–4492

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 7:4055–4062

Herrera-Estrella A, Van Montagu M, Wang K (1990) A bacterial peptide acting as a plant nuclear-targeting signal: the *Agrobacterium* VirD2 directs β-galactosidase into tobacco nuclei. Proc Natl Acad Sci USA, 87: in press

Jen GC, Chilton M-D (1986) The right border region of pTiT37 T-DNA is intrinsically more active than the left border region in promoting T-DNA transformation. Proc Natl Acad Sci USA 83:3895–3899

Joos H, Timmerman B, Van Montagu M, Schell J (1983) Genetic analysis of transfer and stabilization of *Agrobacterium* DNA in plant cells. EMBO J 2:2151–2160

Koncz C, Martini N, Mayerhofer R, Koncz-Kalman Z, Körber H, Redei GP, Schell J (1989) High-frequency T-DNA-mediated gene tagging in plants. Proc Natl Acad Sci USA 86:8467–8471

Lemmers M, De Beuckeleer M, Holsters M, Zambryski P, Depicker A, Hernalsteens J-P, Van Montagu M, Schell J (1980) Internal organization, boundaries and integration of Ti-plasmid DNA in nopaline crown gall tumours. J Mol Biol 144:353–376

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Mooslehner K, Karls U, Harbers K (1990) Retroviral integration sites in transgenic Mov mice frequently map in the vicinity of transcribed DNA regions. J Virol 64:3056–3058

Peralta EG, Hellmiss R, Ream W (1986) *Overdrive*, a T-DNA transmission enhancer on the *A. tumefaciens* tumour-inducing plasmid. EMBO J 5:1137–1142

Simpson RB, Spielmann A, Margossian L, McKnight TD (1986) A disarmed binary vector from *Agrobacterium tumefaciens* functions in *Agrobacterium rhizogenes*. Plant Mol Biol 6:403–415

Spielmann A, Simpson RB (1986) T-DNA structure in transgenic tobacco plants with multiple independent integration sites. Mol Gen Genet 205:34–41

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 318:624–629

Stachel SE, Timmerman B, Zambryski P (1986) Generation of single-stranded T-DNA molecules during the initial stages of T-DNA transfer from *Agrobacterium tumefaciens* to plant cells. Nature 322:706–712

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: requirements for 5′ *virD* gene products. EMBO J 6:857–863

Teeri T, Herrera-Estrella L, Depicker A, Van Montagu M, Palva ET (1986) Identification of plant promoters in situ by T-DNA-mediated transcriptional fusions to the *npt*-II gene. EMBO J 5:1755–1760

Van Lijsebettens M, Inzé D, Van Montagu M, Schell J (1986) Transformed cell clones as a tool to study T-DNA integration mediated by *Agrobacterium tumefaciens*. J Mol Biol 188:129–145

Wallroth M, Gerats AGM, Rogers SG, Fraley RT, Horsch RB (1986) Chromosomal localization of foreign genes in *Petunia hybrida*. Mol Gen Genet 202:6–15

Wang K, Herrera-Estrella L, Van Montagu M, Zambryski P (1984) Right 25-bp terminus sequences of the nopaline T-DNA is essential for and determines direction of DNA transfer from *Agrobacterium* to the plant genome. Cell 38:455–462

Wang K, Stachel S, Timmerman B, Van Montagu M, Zambryski P (1987) Site-specific nick in the T-DNA border sequence following *vir* gene expression in *Agrobacterium*. Science 235:587–591

Yanofsky MF, Porter SG, Young C, Albright LM, Gordon MP, Nester EW (1986) The *virD* operon of *Agrobacterium tumefaciens* encodes a site-specific endonuclease. Cell 47:471–477

Young C, Nester EW (1988) Association of the VirD2 protein with the 5′ end of T strands in *Agrobacterium tumefaciens*. J Bacteriol 170:3367–3374

Communicated by J. Schell