Plant Molecular Biology **31**: 677–681, 1996.

Short communication

Deviating T-DNA transfer from *Agrobacterium tumefaciens* to plants

Eric van der Graaff, Amke den Dulk-Ras and Paul J.J. Hooykaas
Institute of Molecular Plant Sciences, Leiden University, Clusius Laboratory, Wassenaarseweg 64, 2333 AL Leiden, Netherlands

Received 10 January 1996; accepted in revised form 4 March 1996

Key words: *Agrobacterium tumefaciens*, *Arabidopsis thaliana*, T-DNA transfer, transformation

Abstract

We analyzed 29 T-DNA inserts in transgenic *Arabidopsis thaliana* plants for the junction of the right border sequences and the flanking plant DNA. DNA sequencing showed that in most lines the right border sequences transferred had been preserved during integration, corroborating literature data. Surprisingly, in four independent transgenic lines a complete right border repeat was present followed by binary vector sequences. Cloning of two of these T-DNA inserts by plasmid rescue showed that in these lines the transferred DNA consisted of the complete binary vector sequences in addition to the T-region. On the basis of the structure of the transferred DNA we propose that in these lines T-DNA transfer started at the left-border repeat, continued through the vector part, passed the right border repeat, and ended only after reaching again this left-border repeat.

The bacterium *Agrobacterium tumefaciens* transfers a segment of its tumor-inducing (Ti) plasmid to plant cells during tumor formation [for reviews see 4, 15, 20]. This T(transferred) region is flanked by 24 bp T-DNA border repeats in the Ti plasmid. T-DNA transfer is mediated by the *Agrobacterium vir* (virulence) system, which is inducible by plant phenolic compounds. Induction of the *vir* system leads to nicking of the border repeats at a precise site by VirD proteins, during which the VirD2 protein remains covalently bound to the nucleotide at the 5′ end of the nick site. It is thought that the 3′ OH end of the right-border nick forms the start point for DNA synthesis leading to strand displacement in a polar fashion from the right border repeat (RB) to the left-border repeat (LB). This leads eventually to the release of a single-stranded copy of the T-region (T-strand).

Due to the positions of the nick sites, T-strands theoretically embrace only 2 bp of the RB, but 22 bp of the LB [9]. Indeed, when the yeast *Saccharomyces cerevisae* was used as a recipient of T-DNA, T-circles were found containing one intact T-DNA border repeat, the structure of which was compatible with one originating from ligation of these partial border repeats [1]. Analysis of the junction regions between the T-DNA border repeats and flanking DNA after T-DNA integration in plants revealed that often one or both of the two nucleotides of the transferred RB had been preserved. However, the LB was usually deleted possibly because this end lacks protection by the VirD2 protein [7, 8, 15]. The presence of the covalently bound VirD2 protein at the 5′ end of the RB [19] may not only direct the T-strand efficiently to the nucleus [5, 14], but also protect this end from degradation during T-DNA inte-

gration [2]. Indeed, a specific mutation in VirD2 led not only to reduced T-DNA transfer, but also to more frequent deletions of the RB of the T-DNA [16].

The *Agrobacterium* plant vector system allows the transfer of any DNA segment to plant cells, provided that it is surrounded by the 24 bp T-DNA border repeats. To accommodate plant genetic engineering, plant binary vectors were developed in which the T-DNA is present on a small replicon. Recently it was reported that not only the T-DNA, but also vector parts are sometimes transferred from agrobacteria carrying such binary vectors [6]. It is thought that these vector sequences are transferred because of readthrough of the LB during T-DNA processing. However, below we shall provide evidence that an alternative mechanism may also underlie the transfer of vector sequences.

A pool of transgenic *Arabidopsis* plants harboring the activator T-DNA construct pSDM1550 (see Fig. 2) [17] was screened by us for altered phenotypes. Putative lines exhibiting such an altered phenotype were analyzed for linkage of this mutant phenotype with the T-DNA insert. The right ends of the T-DNAs and the plant DNAs flanking these right ends were cloned from these lines using plasmid rescue [12, 17]. Subsequently, the flanking plant DNA was analyzed by sequence analysis using a sequence primer (AS-1: 5′-CCACTGACGTAAGGGATGAC-3′) annealing in the 35S promoter region of the T-region near the RB of the pSDM1550 vector.

From 12 independent transgenic lines 29 T-DNA inserts were isolated and from these the RB/plant DNA fusion sequences were thus characterized. Eight of the isolated plasmids were derived from T-DNA repeat structures [17]. The RB/plant DNA fusion sequences of the remaining 21 T-DNA inserts are presented in Fig. 1. The conservation of the RB sequences after T-DNA integration was fairly precise, leading to only small deletions of T-DNA sequences at this RB as was reported before. In only one case (357H41) a somewhat larger deletion comprising 24 bp of the right part of the T-DNA was encountered.

Surprisingly, in four out of the 29 plasmids

	T-DNA	right border nick site ⇓
pSDM1550		ATCTAGATCTCGA**GGAGGGATATATACCGTTG**
Line #	plasmid #	
17	H	ATCTAGATCTCGA**GGAGGGATATATACCGTTG**
80	B1	ATCTAGATCTgagtgtgccatga
99	H16	ATCTAGATCTCGA**GG**gagctcgc
99	H17	ATCTAGATCTCGA**GG**tatcacat
99	H20	ATCTAGATCTCGA**GG**ccaatcat
99	H35	ATCgctccgcgacgtctgtcgag
234	H	ATCTAGATCTCGA**G**actctaaag
300	H	ATCTAGATCTCGA**GG**aatgtggt
329	H	ATCTAGATCTCGA**GG**gtatcaca
357	H37	ATCTAGATCTCGA**G**aacactggt
357	H41	actgacttttgatggatatattt
367	H30	ATCTAGATCTCGA**G**ttgaattca
404	H5	ATCTAGAcatggatgttcacttt
404	H12	ATCTAGATCTCGAtcggtcctgc
491	H40	ATCcagacctgcactgcacctga
491	P68	ATCTAGATCTCGA**GGAGGGATATATACCGTTG**
491	B119	ATCTAGATCTCGA**G**acactggta
566	H8	ATCTAGATCTCGA**GGAGGGATATATACCGTTG**
566	B25	ATCTAGAgaagagagtcagaaaa
732	H73	ATCTAGATCTCGA**GGAGGGATATATACCGTTG**
732	H76	ATCTAGATCTCGA**GG**tatcacat

Fig. 1. Right border fusion sequences of the rescued plasmids from the gene tagging approach. The nucleotides of the T-region are presented in capitals with the T-DNA right border repeat sequences in bold, while the plant DNA sequences are given in lower case. The sequence of the right border repeat of plasmid pSDM1550 is shown on top for comparison.

(17H, 491P68, 566H8 and 732H73) sequences of the vector part of the binary vector were encountered in the rescued plasmids. Restriction enzyme analysis of these four plasmids showed that they all contained the restriction sites characteristic for the vector part of the binary vector (see fig. 2). Identical restriction fragments were identified in the transgenic plant lines by Southern blot analysis (results not shown) thus excluding contamination during plasmid rescue.

To exclude the possibility that the transgenic plant lines were infected with agrobacteria containing pSDM1550 DNA preparations isolated from these four lines were analyzed by PCR using primers specific for the *virB* region of *A. tumefaciens*. No amplification occurred, whereas control primers showed clear amplification of a segment of the T-DNA, indicating that the lines do not contain detectable bacterial contamination (data not shown).

Subsequently, larger junction clones were isolated from two of the four lines (lines 17 and 732). Analysis of these two clones using restriction enzyme digestions and by DNA sequencing

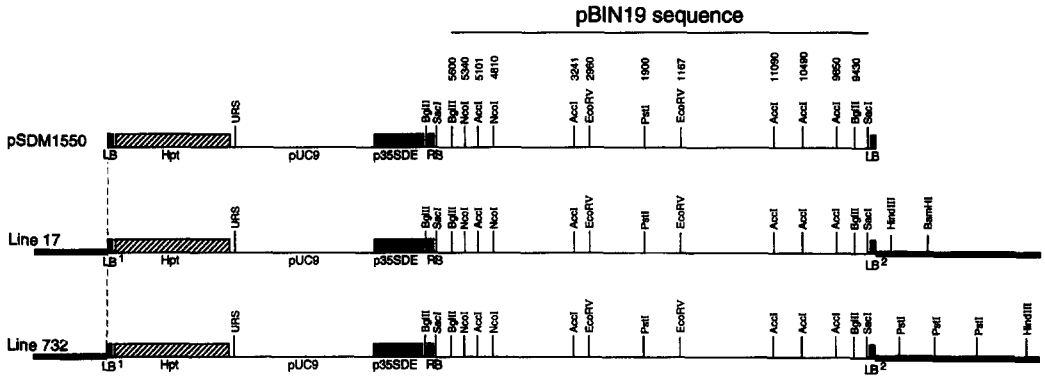

Fig. 2. Schematic map of the pSDM1550 gene tagging vector and the plasmids rescued from the lines 17 and 732 harboring the binary vector sequences. The restriction sites used to determine the presence of binary vector sequences in these rescued plasmids are represented for each line. The corresponding map positions of these restriction sites in the pBIN19 vector [3] on which the pSDM1550 vector is based are given for the pSDM1550 vector. URS, unique restriction sites used for plasmid rescue (*Bam*HI, *Sal*I, *Pst*I, *Sph*I and *Hin*dIII); pUC9, complete pUC9 cloning vector; p35SDE, 35S CaMV promoter with doubled enhancer and AMV leader sequence; Hpt, hygromycin resistance marker; RB, right T-DNA border repeat; LB, left T-DNA border repeat; LB^1, 22 bp of the LB, right from the nick site; LB^2, 2 bp (GG) of the LB, left from the nick site.

showed for both clones that the RB of the T-DNA is flanked by the complete binary vector up to the LB, which in turn is linked to plant DNA (see Fig. 2). Sequence analysis of the LB/plant DNA junction region of these two T-DNA loci (see Fig. 3) showed that only the two GG nucleotides of the LB left from the nick site were present in these T-DNA loci, whereas sequences to the right of this nick site in the LB were derived from plant DNA. The fact that the LB ends of these T-DNAs were fused to plant DNA and that these plant sequences differed for the two lines corroborates the results presented above that these plasmids

```
          Binary vector  left border nick site
                                 ⇓
pSDM1550 GCTTTCGCGAGCTC**GGCAGGATATTCAATT**

Line #

17       GCTTTCGCGAGCTC**GG**gatcgatcggtaacac
732      GCTTTCGCGAGCTC**GG**catttgctatgggtta
```

Fig. 3. Left border fusion sequences of the rescued plasmids harboring the binary vector sequences from line 17 and 732. The nucleotides of the binary vector are presented in capitals with the T-DNA left-border repeat sequences in bold, while the plant DNA sequences are given in lower case. The sequence of the left-border repeat of plasmid pSDM1550 is shown on top for comparison.

were not isolated due to a contamination with the original transformation vector during the plasmid rescue procedure. Moreover, these data indicate that the transfer of the binary vector sequences flanking the RB in these four transgenic lines probably results from T-DNA processing, starting at the LB rather than at the RB followed by skipping of the RB during processing. The occurrence of border skipping during T-strand synthesis has also been detected in *Agrobacterium* following acetosyringone induction [13].

In order to determine whether these four unusual lines had integrated a segment larger than one unit of the complete binary vector, which is theoretically possible after rolling circle replication and skipping of the border repeats, DNA preparations from the lines were analyzed for the presence of a segment embracing the LB by PCR. The primers used (p35Sout: 5′-CGCCAGTCTTTACGGCGAGT and LBpSDM14: 5′-ATCTTGCTCGTCTCGCTGGC) allow the detection of a DNA segment of 500 bp including the LB. In a control experiment about 60 T-DNA insertions were analyzed, yielding 3 PCR positives (results not shown), indicating readthrough over the LB in 5% of the T-DNAs. However, none of the four transgenic lines harboring the binary vector

T-DNA loci showed this 500 bp PCR product, which is indicative of skipping of the LB.

Southern blot analysis of DNA preparations from these transgenic lines showed that none of these lines harbored truncated T-DNA copies besides the long T-DNA (data not shown). Genetic analysis confirmed that the hygromycin resistance trait (left part of the T-DNA) is cosegregating with the right part of the T-DNA and the attached binary vector sequences. Therefore, we favor the hypothesis that the four unusual transgenic lines received a T-strand of which transfer started at the LB leading to transfer of the vector part, continued through the RB into the T-region, and ended again at the LB resulting in transfer of one complete unit of the binary vector linearized at the LB (see Fig. 4). Subsequent plasmid rescue of such T-DNA inserts would result in the rescue of the complete RB with the flanking binary vector sequences as found for four of the transgenic lines in this study. The finding that in the two lines analyzed for the LB/plant DNA junction only the two GG nucleotides left from the left border nick site were present indeed corroborates that T-strand synthesis originated at this LB. During transfer the VirD2 protein probably remained covalently attached to this LB 5′ end, resulting in exact integration of this part of the LB [2].

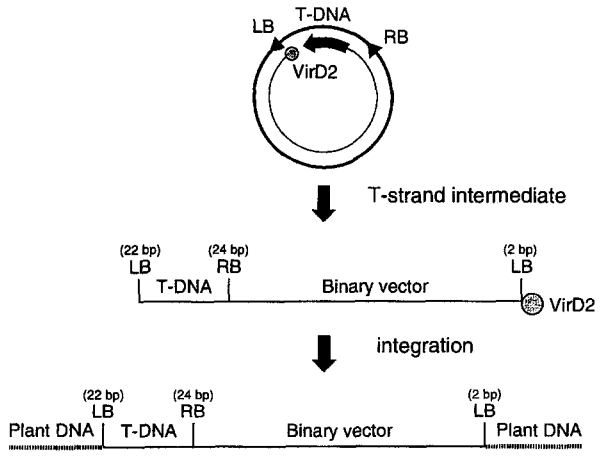

Fig. 4. Model for the transfer and integration of the binary vector sequences into the plant genome for the lines 17, 491, 566 and 732. The number of nucleotides from the LB and RB that are present in the T-strand intermediate and at the T-DNA locus after integration are given.

That T-DNA processing can be initiated at the LB was found long ago, when the activity of the LB and the RB were compared. It was found then that there is no intrinsic difference between the LB and the RB, although the RB is more active due to the presence of overdrive sequences [10, 18]. However, since the overdrive sequence exerts its activity even when it is distal from the RB [18], it is possible that it may stimulate transfer not only from the RB, which is closeby, but also from the distal LB. The effect may depend then on the size of the intervening T-region, and on the DNA topology, i.e. whether the T-region is present on the Ti plasmid or on a small binary vector. Recently, T-DNA transfer starting from the LB from its natural position within the Ti plasmid was detected by insertion of a selection marker near this LB, but outside the T-region [11]. The frequency of LB-mediated transfer was 1–2% of that caused by the RB. In our experiments 4 out of 29 T-DNAs were due to processing starting from the LB and skipping of the RB. It can be assumed that the frequency of LB-mediated transfer was even higher, since it is likely that most of the T-DNA processing will, terminate at the RB in this case. The resulting 'T-strand' (representing the vector part of the binary vector), however, was not selected for in our transgenic plants, and that may be the reason that we did not find single integrated vector parts. If we assume that in our experiments LB-mediated transfer was as frequent as RB-mediated transfer, we can conclude that RB skipping (about 20%) must have been much less rare than LB skipping (about 5%). This may be due to the fact that our experiments were biased by selection for transgenic plants with an altered phenotype. Further experimentation will have to be done to resolve this.

Acknowledgements

This work was supported by the Netherlands Foundation for Biological Research (SLW) and by the European Communities BIOTECH Programme, as part of the Project of Technological Priority 1993–1996.

References

1. Bundock P, den Dulk-Ras A, Beijersbergen A, Hooykaas PJJ: Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. EMBO J 14: 3206–3214 (1995).
2. Dürrenberger F, Crameri A, Hohn B, Koukolíková-Nicola Z: Covalently bound VirD2 protein of *Agrobacterium tumefaciens* protects the T-DNA from exonucleolytic degradation. Proc Natl Acad Sci USA 86: 9154–9158 (1989).
3. Frisch DA, Harris-Haller LW, Yokubaitis NT, Lomas TL, Hardin SH, Hall TC: Complete sequence of the binary vector Bin 19. Plant Mol Biol 27: 405–409 (1995).
4. Hooykaas PJJ, Beijersbergen AGM: The virulence system of *Agrobacterium tumefaciens*. Annu Rev Phytopath 32: 157–179 (1994).
5. Howard EA, Zupan JR, Citovsky V, Zambryski PC: The VirD2 protein of *Agrobacterium tumefaciens* contains a C-terminal bipartite nuclear localization signal: implications for nuclear uptake of DNA in plant cells. Cell 68: 109–118 (1992).
6. Martineau B, Voelker TA, Sanders RA: On defining T-DNA. Plant Cell 6: 1032–1033 (1994).
7. Matsumoto S, Ito Y, Hosoi T, Takahashi Y, Machida Y: Integration of *Agrobacterium* T-DNA into a tobacco chromosome: possible involvement of DNA homology between T-DNA and plant DNA. Mol Gen Genet 224: 309–316 (1990).
8. Mayerhofer R, Koncz-Kalman Z, Nawrath C, Bakkeren G, Crameri A, Angelis K, Redei GP, Schell J, Hohn B, Koncz C: T-DNA integration: a mode of illegitimate recombination in plants. EMBO J 10: 697–704 (1991).
9. Pansegrau W, Lanka E: Common sequence motifs in DNA relaxases and nick regions from a variety of DNA transfer systems. Nucl Acids Res 19: 3455 (1991).
10. Peralta EG, Hellmiss R, Ream W: Overdrive, a T-DNA transmission enhancer on the *Agrobacterium tumefaciens* tumour-inducing plasmid. EMBO J 5: 1137–1142 (1986).
11. Ramanathan V, Veluthambi K: Transfer of non T-DNA portions of the *Agrobacterium tumefaciens* Ti plasmid pTiA6 from the left terminus of T_L-DNA. Plant Mol Biol 28: 1149–1154 (1995).
12. Rommens CMT, Rudenko GN, Dijkwel PP, Van Haaren MJJ, Ouwerkerk PBF, Blok KM, Nijkamp HJJ, Hille J: Characterization of the *Ac/Ds* behaviour in transgenic tomato plants using plasmid rescue. Plant Mol Biol 20: 61–70 (1992).
13. Stachel SE, Timmerman B, Zambryski P: Activation of *Agrobacterium tumefaciens vir* gene expression generates multiple single-stranded T-strand molecules from the pTiA6 T-region: requirement for 5′ *virD* gene products. EMBO J 6: 857–863 (1987).
14. Tinland B, Koukolíková-Nicola Z, Hall MN, Hohn B: The T-DNA linked VirD2 protein contains two distinct functional nuclear localization signals. Proc Natl Acad Sci USA 89: 7442–7446 (1992).
15. Tinland B, Hohn B: Recombination between prokaryotic and eukaryotic DNA: integration of *Agrobacterium tumefaciens* T-DNA into the plant genome. In: Setlow JK (ed) Genetic Engineering vol 17, pp. 209–229 Plenum Press, New York (1995).
16. Tinland B, Schoumacher F, Gloeckler V, Bravo-Angel AM, Hohn B: The *Agrobacterium tumefaciens* virulence D2 protein is responsible for precise integration of T-DNA into the plant genome. EMBO J 14: 3585–3595 (1995).
17. van der Graaff E, Hooykaas PJJ: Improvements in the transformation of *Arabidopsis thaliana* C24 leaf discs by *Agrobacterium tumefaciens*. Plant Cell Rep, in press (1996).
18. van Haaren MJJ, Sedee NJA, Schilperoort RA, Hooykaas PJJ: Overdrive is a T-region transfer enhancer which stimulates T-strand production in *Agrobacterium tumefaciens*. Nucl Acids Res 15: 8983–8997 (1987).
19. Ward ER, Barnes WM: VirD2 protein of *Agrobacterium tumefaciens* very tightly linked to the 5′ end of T-strand DNA. Science 242: 927–930 (1988).
20. Zambryski PC: Chronicles from the *Agrobacterium* plant cell DNA transfer story. Annu Rev Plant Physiol Plant Mol Biol 43: 465–490 (1992).