Targeted recombination in plants using Agrobacterium coincides with additional rearrangements at the target locus

Risseeuw, E.; Offringa, R.; Franke-Van Dijk, M.E.I.; Hooykaas, P.J.J.

The Plant Journal (1995) 7(1), 109–119

Targeted recombination in plants using *Agrobacterium* coincides with additional rearrangements at the target locus

Eddy Risseeuw*, Remko Offringa, Marry E.I. Franke-van Dijk and Paul J.J. Hooykaas
Leiden University, Clusius Laboratories, Institute of Molecular Plant Sciences, Wassenaarseweg 64, 2333 AL Leiden, The Netherlands

Summary

The use of *Agrobacterium* for gene targeting in plants has been investigated. Leaf protoplasts of five transgenic tobacco lines, containing a T-DNA insertion with a defective *npt-II* gene at different positions in the plant genome, were transformed via *Agrobacterium* with a T-DNA containing a *npt-II* repair gene. After selection for kanamycin resistance and PCR analysis six recombinants were derived from four of the target lines. The recombination frequencies were similar for the different target lines with one recombinant from approximately 3×10^5 transformants. Apparently gene targeting is more or less independent of the location of the target construct in the plant genome. Molecular analysis revealed that gene targeting had occurred in five of the six recombinant lines. However precise recombination had occurred in only one line, while in the other four lines restoration of the *npt-II* gene was accompanied by a deletion of part of the target locus. The sixth recombinant line showed restoration of the *npt-II* gene of the incoming T-DNA construct which was inserted in the plant genome at a position closely linked to the target locus. The different recombination products favour a model in which recombination is via gene conversion followed by reintegration of the synthesized DNA via homologous or illegitimate recombination rather than a reciprocal exchange of DNA between two cross-overs.

Introduction

Gene targeting, that is the directed integration of introduced DNA into the genome via homologous recombination, can be a valuable tool for many molecular genetic studies. Genes of interest can be mutated or inactivated and new sequences can be introduced at defined loci of the genome. In prokaryotes and some lower eukaryotes, where introduced DNA is predominantly integrated via homologous recombination, gene targeting has become a standard technique to obtain specific mutants. In higher eukaryotes, however, introduced DNA is predominantly integrated at random positions. Still, for some mammalian systems, gene targeting has now become a standard technique with targeting frequencies of more than 1% having been reported. Targeting events have been reported to occur in plant cells as well at lower frequencies ($< 5\times10^{-4}$) (for review, see Offringa *et al.*, 1992; Ohl *et al.*, 1994). These results have been obtained both after direct DNA transfer (Halfter *et al.*, 1992; Paszkowski *et al.*, 1988) and after *Agrobacterium*-mediated T-DNA transfer (Lee *et al.*, 1990; Offringa *et al.*, 1993) to plant cells.

Previously we reported the targeted correction of a target locus containing a defective kanamycin-resistance (*npt-II*) gene in tobacco cells after homologous recombination with an incoming repair T-DNA (Offringa *et al.*, 1990). More recently, we showed that the opposite event, i.e. correction of the defective *npt-II* gene at the repair T-DNA after homologous recombination with the target locus, also occurs (Offringa, 1992; Offringa *et al.*, 1993).

Our main objective was to study which factors determine the low targeting frequencies in plants. Although many factors are known to influence the targeting frequency, several groups have attributed large differences in the targeting frequency to a different position in the genome (Halfter *et al.*, 1992; Paszkowski *et al.*, 1988). In this report we tested this hypothesis by analysing the targeting frequencies for a particular gene construct located at five different positions in the tobacco genome. An additional goal was to select a target line with a high targeting frequency which might be interesting for further experiments. The recombinant lines we obtained were analysed in detail by PCR and Southern blot. This revealed that most recombination events were accompanied by additional rearrangments at the target locus. On the basis of the results a model for the mechanism of recombination is postulated.

Results

Experimental design

The T-DNA of the binary vector pSDM304 containing a defective kanamycin resistance (*npt-II*) gene next to a hygromycin-resistance (*hpt*) gene was stably integrated in the genome of *Nicotiana tabacum* SR1 after transformation of leaf-discs with *Agrobacterium tumefaciens* (Figure 1). Hygromycin-resistant plant lines were obtained and lines with single copy T-DNA insertions were selected by inverse PCR and Southern blot analysis (Does *et al.*, 1991). Because of the 3′ deletion in the *npt-II* gene, the transformed lines

Received 11 August 1994; revised 10 October; accepted 14 October 1994.
*For correspondence (fax 31 71 274999).

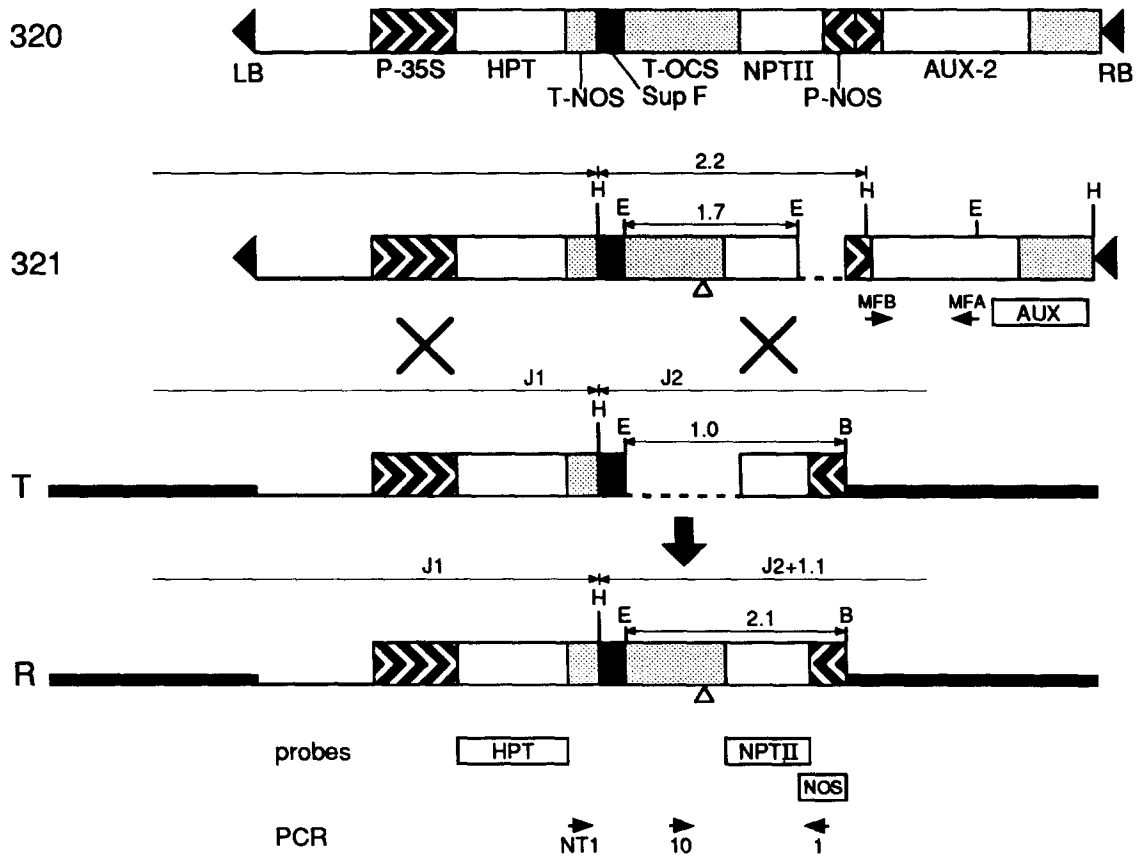

Figure 1. Experimental design for the detection of gene targeting.
Transgenic *Nicotiana tabacum* SR1 plants were generated by transformation with T-DNA construct pSDM304 containing a *hpt* gene and a *npt-II* gene which was defective at the 3' side. These target lines (T) were retransformed with the replacement vector pSDM321. The predicted situation after gene targeting (R) is indicated. Construct pSDM320 was used for estimation of the transformation frequency in the targeting experiments. Genomic DNA fragments obtained after digestion with EcoRI/*Bcl*I and with *Hin*dIII, and hybridizing with the *hpt* and *npt-II* probes are indicated. Primers used for the PCR analysis are indicated with arrows.
P-35S, promoter of the 35S transcript of cauliflower mosaic virus; HPT/NPTII, hygromycin/neomycin phosphotransferase coding region; P-NOS, T-NOS, promoter, terminator of the nopaline synthase gene; T-OCS, terminator of the octopine synthase gene; AUX-2, the pTi *aux-2* gene; SupF, *E. coli* supressor gene F; LB, RB, left and right border T-DNA repeats respectively; △, marker deletion; H, *Hin*dIII; E, *Eco*RI; B, *Bcl*I; J1, J2, junction fragments.

were kanamycin sensitive. These primary target lines were propagated and then transformed once more by an *Agrobacterium* strain with the replacement vector pSDM321 containing a *npt-II* gene lacking the promoter region and the first 12 codons of the open reading frame. At the left arm a 3.5 kb segment and at the right arm 0.6 kb of homology with the target locus was present within the T-DNA of pSDM321. Homologous recombination by double cross-over or conversion between target and repair construct will result in an intact *npt-II* gene leading to kanamycin resistance. Initial experiments showed a high background of Km^r calli in which the *npt-II* gene in the replacement vector was activated by fusion to a plant gene or the *hpt* gene on the vector (De Groot *et al.*, 1994). To reduce this background a non-homologous *aux-2* gene from the *A. tumefaciens* T-region was cloned in front of the defective *npt-II* gene in construct pSDM321 (Offringa *et al.*, 1993). In addition, a marker deletion was introduced on the 3' side of the *npt-II* coding region to simplify PCR analysis. To detect even recombination events occurring at a low frequency, we used co-cultivation of tobacco leaf protoplasts with *A. tumefaciens* to obtain large numbers of transformants. The kanamycin resistant (Km^r) calli obtained were checked for the presence of a recombination product by PCR using primers 1 and 10 located in the deleted parts of the *npt-II* gene (Figure 1). This leads to the amplification of a 1 kb fragment in recombinant lines. The presence of the marker deletion in the PCR fragment can exclude contamination with any plasmid construct during transformation.

Targeting experiments

Leaf protoplasts from five selected tobacco lines, each containing a single target T-DNA, were transformed with replacement vector pSDM321 and Km^r calli were selected.

Table 1. Targeting to different loci of the tobacco genome via *A. tumefaciens* with the targeting vector pSDM321

Experiment	Target line	Number of protoplasts	Transformed calli[a]	Km^r calli	PCR pos. calli[b]	Recombination frequency[c]
III	304.4	8×10^6	4.9×10^5	441	1/407	2.2×10^{-6}
	304.5	8×10^6	4.9×10^5	485	2/417	4.8×10^{-6}
	304.12	8×10^6	4.4×10^5	498	1/420	2.7×10^{-6}
IV	304.32	12×10^6	3.8×10^5	239	2/212	6.3×10^{-6}
	304.35	10×10^6	2.6×10^5	215	0/211	-
Total		46×10^6	2.1×10^6	1878	6/1667	3.2×10^{-6}

[a]The predicted total number of transformed calli with pSDM321, as calculated from the number of Km^r calli obtained with the control construct pSDM320.
[b] Number of tested calli which were PCR positive with primers 1 and 10.
[c] Frequency related to the transformation frequency with pSDM320.

The number of Km^r calli obtained with this construct was about 1000-fold lower than that found after transformation with the T-DNA construct pSDM320 containing the intact *npt-II* gene (Table 1 and Figure1). No Km^r calli were obtained after transformation with an *Agrobacterium* strain lacking a binary vector. Pools of 25 Km^r calli obtained from the targeting experiments were screened for the presence of a recombination product by PCR using primers 1 and 10 (Figure 1). Individual calli from positive pools were further examined and thus single positive calli were identified. In this way six recombinant calli were obtained from targeting experiments with four target lines. On average one recombinant was found in 3.1×10^5 transformants. None of the target lines tested showed a significantly higher recombination frequency.

Molecular analysis of putative recombinants

The putative recombinant lines were further analysed by Southern blotting. Genomic DNA was digested with *Eco*RI/*Bcl*I and hybridized with the *npt-II* probe to screen for the presence of the recombinant 2.1 kb fragment and absence of the original 1.0 kb fragment (Figure 2a). The expected fragment shift was found in four lines: R_1, R_2, R_4 and R_6. In line R_3 both the recombinant and the original fragment were present. In line R_5 the original 1.0 kb fragment was absent, but this had shifted to a fragment much larger than the expected 2.1 kb. Besides the recombinant fragment, a strongly hybridizing 1.7 kb DNA fragment was found in all recombinant lines, which represents the 5′ deleted *npt-II* gene from the repair T-DNA and is due to the random integration of the repair construct at various loci in the genome. Digestion with *Hin*dIII and hybridization with the *npt-II* probe reveals the junction fragment J_2 (Figure1). After homologous recombination this fragment should be enlarged by 1.1 kb. Indeed the original J_2 fragment of the target locus had disappeared and a new DNA fragment of the expected larger size was detected in all the recombinant lines with the exception of line R_3 (Figure 2b). No additional alterations were found on the right arm of the restored target locus. The repair constructs which were randomly integrated revealed an internal fragment of 2.2 kb. In line R_3, the J_2 junction fragment of the orginal target locus was present as well as the expected 1.1 kb larger recombinant fragment. Therefore, it is most likely that in line R_3 the *npt-II* gene of the incoming T-DNA of pSDM321 was restored via gene conversion and integrated elsewhere in the genome. A similar event was described previously (Offringa *et al.*, 1993).

Rearrangements at the targeted loci

The left arm was checked for possible rearrangements caused by the recombination process. For detection of the junction fragment J_1 (Figure 1), the genomic blot in Figure 2b was rehybridized with the *hpt* probe (Figure 2c). This junction fragment J_1 should remain unchanged after correct and precise targeting. However in recombinant lines R_2 and R_4 the J_1 fragment had disappeared. Apparently restoration of the *npt-II* gene at the right arm of the target locus in these lines coincided with rearrangements at the left arm of the target locus. In the other recombinant lines it was difficult to establish unequivocally whether the J_1 fragment was unaltered due to the appearance of many cross-hybridizing bands in the blot, which are due to the random integration of many copies of the targeting vector. Therefore, more restriction sites were mapped at both sides of the target T-DNA in both the target lines and the recombinants. Cross-hybridization with the randomly integrated targeting vectors was prevented by using the *nos* probe which only hybridized with the target locus (Figure 1). An example of the Southern analysis is shown for line R_2 in comparison with the original target line 304.5 (Figure 3). Digestions with *Sph*I and *Nco*I, and subsequently with combinations of *Nco*I and *Kpn*I or *Nco*I and *Xba*I were used to analyse the right arms of the T-DNA. As can be

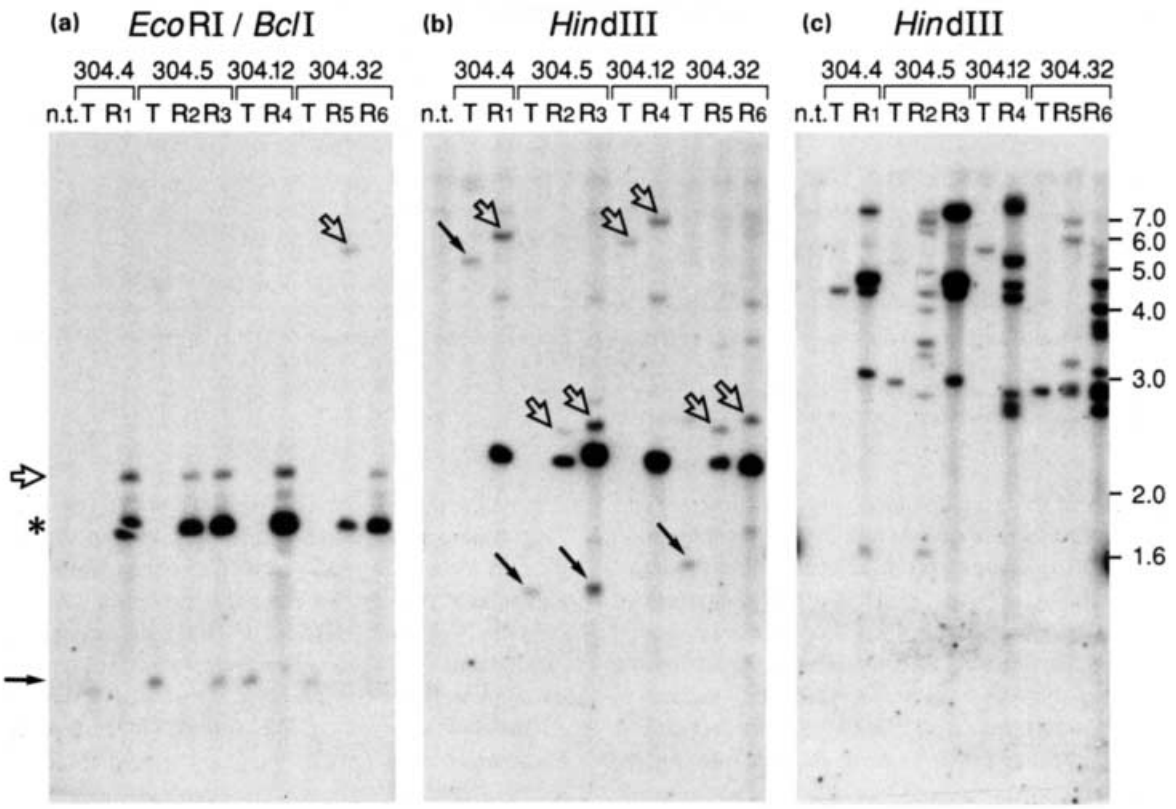

Figure 2. Southern blot analysis of the recombinant lines.
(a) Genomic DNA from the target lines and the recombinant lines digested with *Eco*RI/*Bcl*I and hybridized with the *npt-II* probe. The internal 1.0 kb fragment of the original target locus (→) was detected in the original target lines (T). This fragment showed the expected 1.1 kb shift to a 2.1 kb fragment in the recombinant lines (⇒). R_5 showed a much larger, about 5 kb large recombinant fragment instead of the 2.1 kb fragment. In R_3 both the recombinant 2.1 kb fragment and the original 1 kb fragment were present. Several randomly integrated targeting vectors were present in all recombinant lines, indicated by a strong hybridizing 1.7 kb fragment (*). Other hybridizing fragments were due to aspecific binding or partial digestion. Stochiometric differences between lanes are caused by the slightly different amounts of DNA loaded.
(b) Genomic DNA digested with *Hin*dIII and hybridized with the *npt-II* probe. The *npt-II* probe revealed the junction fragments J_2 of the original target loci (→) which showed the expected 1.1 kb shift in the recombinant lines (⇒). Randomly integrated targeting vectors were identified by the presence of an internal 2.2 kb fragment (*).
(c) Southern blot from (b) hybridized with the *hpt* probe.
n.t., not transformed.

seen in Figure 3 for the right arms of lines R_2 and 304.5 identical fragments were obtained for all of these digestions. The *Eco*RI digestion revealed the expected 1.1 kb shift of the target band in recombinant line R_2. Thus, besides restoration of the defective *npt-II* gene, no alterations were found at the right arm of line R_2. Restriction sites on the left arm were mapped with respect to the *Xba*I and *Bcl*I sites located to the right of the *nos* promoter. In line R_2, none of the digestions with these enzymes resulted in fragments which showed the expected 1.1 kb shift. In fact the restriction map of the left arm of line R_2 did not match with the left arm of the original target locus at all. The results of the mapping of all recombinant lines are summarized in Figure 4. The targeted recombinant lines did not show rearrangements at the right arm of the target locus besides restoration of the *npt-II* gene. However, at the left arm not only in line R_2 but also in R_1, R_4 and R_5 a restriction pattern was found which was different compared with that of the target line. In lines R_1 and R_4 two restriction sites (*Xba*I *Kpn*I and *Bcl*I *Xba*I, respectively) could be mapped on the left arm. These turned out to be the same distance apart in the recombinant line and the original target line. The distance between these sites and the unique *Bcl*I site of target T-DNA revealed a deletion in R_1 and R_4 of 3.0 kb and 2.1 kb, respectively, as compared with the target locus. Recombinant R_6 was the only line which showed an unaltered pattern at the left arm of the target locus and was therefore the result of a perfect gene targeting event.

In R_2 and R_5 no similarities of the restriction patterns could be found between the left arm of the restored locus and the original target locus or the targeting vector pSDM321. The *Bcl*I fragment in R_5 hybridizing with the *nos* probe, had the same size as the hybridizing fragment

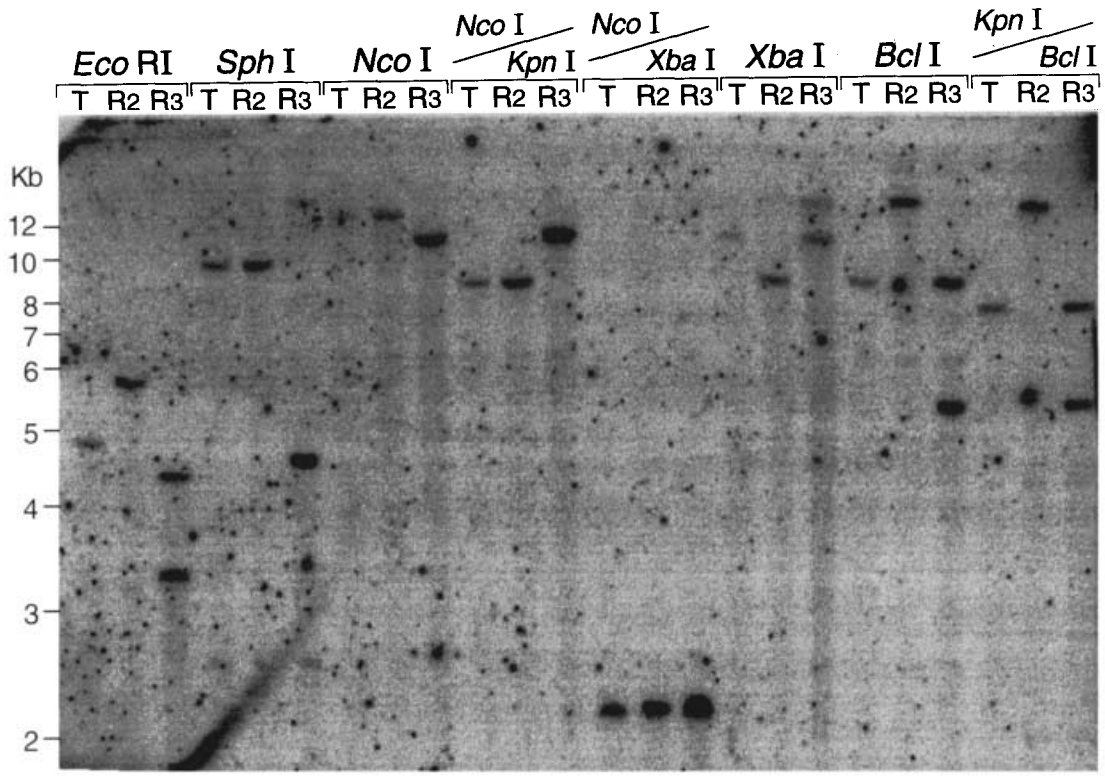

Figure 3. Southern analysis of the recombinant lines R_2 and R_3 with the target line 304.5 (T).
Restriction sites were mapped on the right arm (digestions *Eco*RI up to *Nco*I/*Xba*I) and on the left arm (digestions *Xba*I up to *Kpn*I/*Bcl*I) of the target locus. After exposure of the Southern blot to the Phosphorimager, the picture was generated via Image Quant software.

obtained after digestion with *Bcl*I/*Eco*RI. This was unexpected since the *supF* gene was bounded by two *Eco*RI sites and two other sites were present in the *hpt* gene (data not shown). Thus it appeared that in R_5 the *hpt* and *supF* sequences had been removed from the target locus.

Correction of the targeting vector coinciding with a rearrangement at the target locus

Southern analysis of R_3 revealed that the *npt-II* gene at the incoming targeting vector had been corrected (see above). Further restriction enzyme analysis showed that the left arms of the target locus of 304.5 and of the corrected targeting vector were different (Figures 3 and 4). At the right arm of the targeting vector, the *aux-2* gene was exchanged with the *nos* promoter and a part of the genomic sequence next to the target T-DNA. The position of the *Hin*dIII and *Xba*I sites bordering the right end of the T-DNA was resolved (Figures 2b and 3); they were at identical positions in the original target locus 304.5 and recombinant R_2 indicating that this segment of DNA had been copied from the target locus towards the targeting vector pSDM321. Restriction sites mapping further right from the *Xba*I site on the original target locus were not found in the repaired targeting vector in line R_3. Remarkably, however in line R_3, no difference was found between the right arm of the restored targeting vector and that of the target locus over a length of more than 10 kb. Southern analysis revealed identical *Sph*I and *Nco*I fragments hybridizing with the *nos* probe. Two fragments were found after digestion with *Eco*RI one being 1.1 kb larger than the other. Thus the position of the *Eco*RI site to the right of the T-DNAs was the same for the target locus and the recombinant line R_3, but different compared with that in the original target line 304.5 and recombinant line R_2 (Figure 3). The repair of the incoming targeting vector pSDM321 apparently caused a rearrangement at the right end of the target locus, resulting eventually in a line R_3 with two T-DNAs with identical DNA segments bordering each of the T-DNAs.

Progeny analysis

The recombinant lines were both selfed and back-crossed with non-transgenic SR1 plants. F_1 seedlings were tested for kanamycin resistance (Table 2). Germination of the seeds was not affected by kanamycin. The presence of the recombinant locus was confirmed in 30 Km^r plants from each back-cross by PCR using primers 1 and 10, excluding

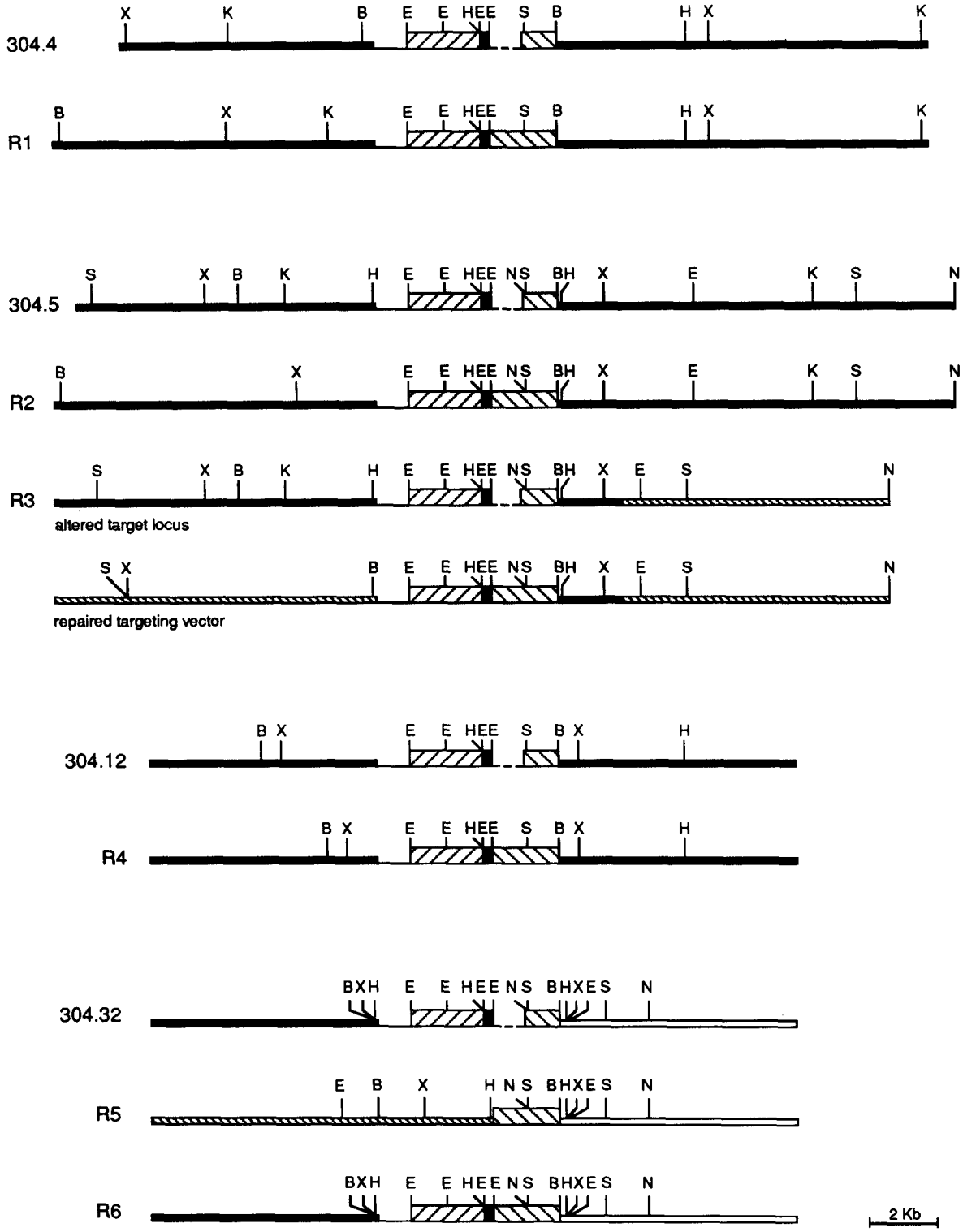

Figure 4. Restriction maps of recombinant target loci in comparison with the original target loci.
X, *Xba*I; K, *Kpn*I; B, *Bcl*I; E,: *Eco*RI; H,: *Hin*dIII; S,: *Sph*I; N, *Nco*I.

Table 2. Progeny analysis of the recombinant lines

Recombinant line	Back-cross with SR1			Selfed		
	Km^r	Km^s	ratio[a]	Km^r	Km^s	ratio[a]
R_1	62	65	1:1	113	31	3:1
R_2	85	48	A	117	46	3:1
R_3	96	20	3:1	85	2	15:1
R_4	56	127	A	32	105	A
R_5	33	49	1:1	52	22	3:1
R_6	55	54	1:1	89	17	3:1

[a] The ratio of kanamycin resistant:sensitive seedlings was compared with the expected ratio in the chi-square test for goodness of fit. The indicated ratios fit at the 1% probability level ($\chi^2<6.63$).
A, aberrant segregation.

the possibility that kanamycin resistance was caused by one of the randomly integrated T-DNAs (Offringa *et al.*, 1990). It was already established by Southern blot analysis that not more than one recombinant *npt-II* gene was present in each recombinant line. Primers MFA and MFB in the *aux-2* gene were used to detect the presence of randomly integrated target vectors. No Km^r F_1 plants were found which had lost all randomly integrated targeting vectors. Three recombinant lines gave aberrant segregation ratios. Line R_4 showed a reduced number of Km^r seedlings, than was expected on the basis of one dominant locus, while R_2 gave the opposite result in the back-cross with SR1. Line R_3 resulted in a segregation ratio that was indicative of the presence of two unlinked loci which is in contradiction with the Southern blot analysis. Moreover, in R_3 no segregation between the target T-DNA and the restored targeting vector was observed by PCR in both the 30 Km^r F_1 plants and 30 randomly picked F_1 plants of the back-cross with SR1 (primers NT1 and 1, Figure 1). This indicated that the target T-DNA and the restored targeting vector are closely linked and that also the target T-DNA segregates for the presence of two unlinked loci.

Discussion

Gene targeting experiments were carried out with five tobacco lines, each containing a single copy of a defective *npt-II* gene. Although no crosses between these lines were performed, Southern analysis showed that the border fragments of the lines were different (Figures 2 and 3) and that at least at the molecular level the target T-DNAs were located at different genomic positions. Restoration of the *npt-II* gene via homologous recombination was detected in four target lines (Table 1). A total number of six recombinant lines were selected. The targeting frequencies were similar in the different lines, with an average of 3.2×10^{-6} recombinants per transformant. This frequency is reminiscent of the earlier reported frequencies using the same model system but a different target line (Ofringa *et al.*, 1990, 1993).

These results indicate that gene targeting can be reproducibly obtained in plants, for different genomic positions with a similar frequency, when using *Agrobacterium*-mediated DNA transfer. Apparently, the position of the target T-DNA in the plant genome was not the limiting factor in our experiments. However, in the experiments of Paszkowski *et al.* (1988) and Halfter *et al.* (1992) homologous recombination with the artificial target locus was only found in a few of the target lines tested (eight events in two out of five lines and four events in three out of 17 lines, respectively). A significant difference with our experiments is that the target loci in our lines consisted of a single intact T-DNA insert, whereas lines of the other two groups were not selected for single plasmid or T-DNA inserts. Multiple inserts might induce silencing of the corrected selection marker, and thus inhibit selection of the recombinant cell (Fujiwara *et al.*, 1993; Matzke *et al.*, 1993). The selection of single copy lines may only partly explain the observed uniformity in gene targeting frequencies. *Agrobacterium* T-DNA is believed to be inserted in regions that are transcriptionally active or have open chromatin structure (Gheysen *et al.*, 1991; Matsumoto *et al.*, 1990; Mayerhofer *et al.*, 1991). Thus, the use of *Agrobacterium* T-DNA transfer to make the target lines may have selected for target loci at genomic positions that are equally accessible for gene targeting.

Until now no significantly higher recombination frequencies than the ones reported in this paper have been published for plants. A reason for the low targeting frequency in our experiments may be the limited homology between target locus and the targeting vector and the large deletion at the target locus which had to be repaired. The size of homology appeared to be an important factor for the recombination frequency in animal cells (Deng and Capecchi, 1992; Hasty *et al.*, 1991). Non-homology was of less importance when both homologous arms of the replacement type vector were of sufficient length (Offringa, 1992; Thomas *et al.*, 1992). A big difference of the arm lengths decreased the recombination frequency (Thomas

Table 3. Summary of the recombinant lines

Experiment	Target line	Recombinant	Correction of:	Rearrangements at target locus:	No. random integrated targeting vectors
III	304.4	R_1	Target locus	At left arm	8
	304.5	R_2	Target locus	At left arm	9
		R_3	T-DNA	At right arm	7
IV	304.12	R_4	Target locus	At left arm	9
		R_5	Target locus	At left arm	6
	304.32	R_6	Target locus	No	9
	304.35	No			

et al., 1992). In our targeting vector the homology was not equally divided over the two arms. The right arm had a length of only 610 bp which might have been a limiting factor for the recombination frequency.

Mechanisms for recombination

A summary of the recombinant lines is shown in Table 3. In five out of six lines the *npt-II* gene at the target locus was restored. Only one of the lines showed a perfect targeting event while in the other four recombinant lines rearrangements were found in the left arm of the target locus (Figure 4). A comparison of the restriction pattern of R_1 and R_4 with their target lines indicated that 3.0 kb and 2.1 kb, respectively, had been deleted. In contrast with R_1 and R_4, the left arms of R_2 and R_5 did not show similarities with the original target line at all. For R_5 it was shown that even sequences within the segment of homology had been removed including the *supF* and the *hpt* genes. Therefore, it is important that both junction fragments of the target locus are checked before assessing the precision of targeting events.

Homologous recombination at the right short arm of the targeting vector is required for the repair of the *npt-II* gene. The rearrangements in the four recombinants can therefore be explained by non-homologous recombination of the left part of the targeting vector with a chromosomal region left of the target T-DNA resulting in a deletion ranging from 2 to more than 7 kb (Figure 5a). Selection for a restored *npt-II* gene located at the right end of the target locus prevents a deletion at this side of the target locus. The fact that recombination of the long arm did not occur within the 3.5 kb of homology suggests recombination via gene conversion using the targeting vector as a template, rather than a reciprocal recombination via two cross-overs. The deletion of a part of the target T-DNA in R_5, would indicate that the conversion was not completed and followed by a reintegration outside of the target T-DNA. Similar deletions were found in homologous recombination studies with mammalian cells using a replacement vector with homology at one side of the construct (Berinstein *et al.*, 1992). At the target locus recombination of the non-homologous arm was observed in regions of non-homology resulting in a deletion. In some recombinant lines the same was also true for the homologous arm.

It may be that the targeting vector was first randomly integrated, followed by an intra or interchromosomal DNA conversion towards the target locus (Figure 5b). Rearrangements at the left arm of the target locus could then be the result of an insertion or substitution via conversion of chromosomal sequences flanking the targeting vector pSDM321. The fact that no rearrangements were found at the right side can be explained by the assumption that a rightwards conversion would either lead to a precise repair of the *npt-II* gene or to integration of *aux-2* sequences in the place of the *nos* promoter. The latter event will not be selected. The multiple random insertions of the targeting vector in all recombinant lines could have increased the chance of intra genomic recombination.

The copying of DNA fragments in recombinant line R_3 confirms that recombination via DNA conversion does occur (Figure 5c and d). In line R_3 the gene conversion was directed towards the targeting vector using the target locus as a template. Via end extension repair also chromosomal sequences right of the target T-DNA were copied. A deletion at the right arm could have been introduced during gene conversion which was conserved at both strands. Subsequently, the extended targeting vector was integrated close to the target locus (Figure 5c). Gene conversion towards a random integrated targeting vector would also be a possibility. In that case the gene conversion was bidirectional. First the target locus up to the *Xba*I site on the right arm was used as a template for conversion towards the targeting vector locus, followed by a switch of templates by which a part of the right arm of this locus was converted towards the target locus (Figure 5d). End extension repair and switching of templates of the same targeting vector pSDM321 was reported previously (Offringa, 1992). Moreover it was shown that the target locus and the restored targeting vector were linked in one recombinant line (Offringa *et al.*, 1993) while in another recombinant no linkage was observed (Offringa, 1992). End

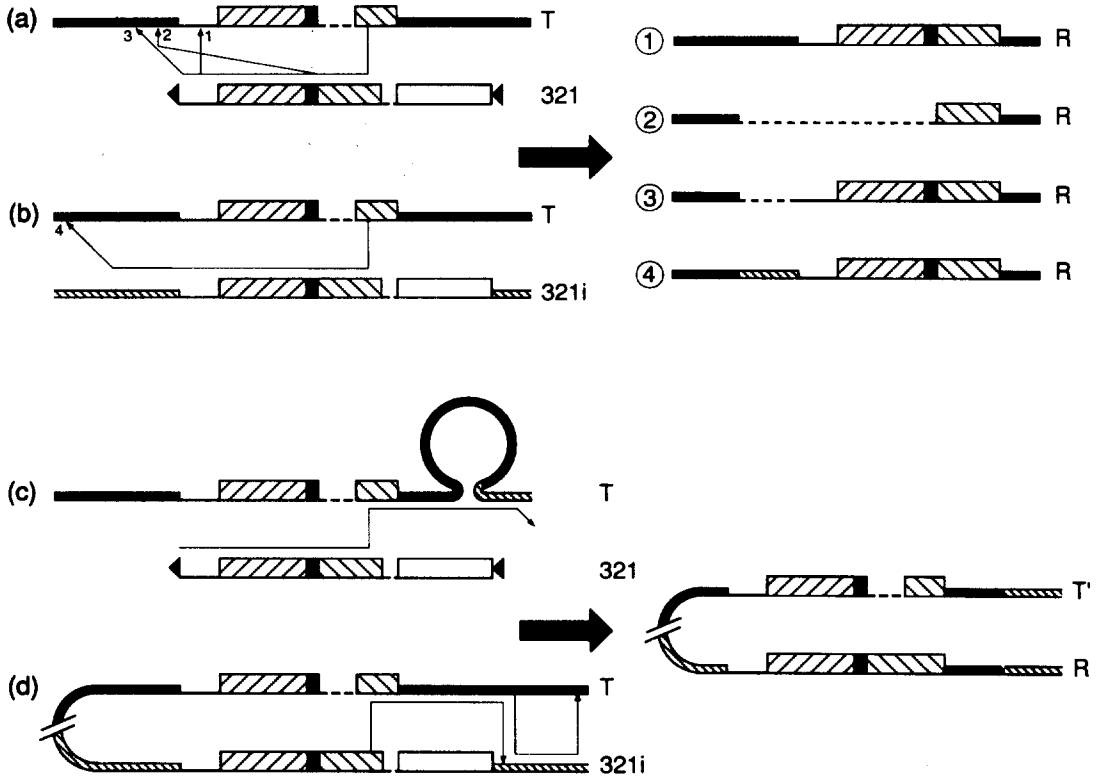

Figure 5. Mechanisms for recombination in plants.
At the target locus, the *npt-I* gene is restored via gene conversion towards the target locus using either the T-DNA from pSDM321 as a template (a) or the targeting vector after integration (b). The length of the conversion and the site of reïntegration result in deletions or substitutions. Gene conversion towards the targeting vector (in R_3) could have occurred either before (c) or after (d) integration. In both situations the targeting vector integrates close to the target locus.

extension repair of targeting vectors was also reported in mammalian cells where this was found with a high frequency (Adair *et al.*, 1989; Aratani *et al.*, 1992). Extension was observed over a length of at least 4.6 kb.

Rearrangements and deletions were also found after illegitimate recombination of the T-DNA into the plant genome (for review, see Koncz *et al.*, 1994). Analysis of T-DNA integration sites revealed that usually several tens of base pairs of the genomic DNA were exchanged by the T-DNA. Rearrangements found at the junctions had a similar size. This is in contrast with the size of the deletions and rearrangements of the recombinant lines described here. In case one of the ends of the T-strand integrates via homologous recombination, illegitimate recombination of the other end seems to happen at a more distant site at the genomic DNA resulting in a larger deletion.

We were not able to determine whether recombination had occurred before or after random integration of the targeting vector. The integration of many extra copies of the targeting vector made a more detailed analysis very difficult. It was striking that all recombinant lines contained a high number of non-targeted integrations of the targeting vector but it is hard to say whether these vectors had contributed to the imprecise recombination events. Analysis of more recombination events will be required to determine whether imprecise recombination is a problem for gene targeting in plants and a cause of the low targeting frequency.

Experimental procedures

Plasmid constructs

Construction of plasmids pSDM320 and pSDM321 was described previously (Offringa *et al.*, 1993). The target construct pSDM304 was different from pSDM104 in that the point mutation in the *npt-II* coding region was restored (Yenofsky *et al.*, 1990). All plasmids were introduced into *Agrobacterium tumefaciens* strain C58C1 harbouring the disarmed Ti plasmid pGV2260, resulting in strains SDM320, 321 and 304.

Plant transformations

The target lines were obtained by co-cultivation of tobacco leaf discs (*Nicotiana tabacum* cv. Petit Havana line SR1) with the *Agrobacterium* strain SDM304 and were selected for hygromycin resistance (Horsch *et al.*, 1985). Subsequently, leaf protoplasts of the target lines were enzymatically isolated and co-cultivated with strains SDM320 and SDM321. Calli were selected in liquid medium

containing up to 100 mg^{-1} kanamycin. Protoplast isolation, co-cultivation and selection of calli were described previously (Offringa *et al.*, 1990).

PCR-analysis

The polymerase chain reactions were performed in a Perkin-Elmer Thermocycler 480. For the screening of recombinant lines, DNA extraction was performed on pools of 25 pieces (100 mg each) of callus by grinding in liquid nitrogen followed by phenol extraction and ethanol precipitation (Mettler, 1987). One microgram of genomic DNA was used in the PCR with 3 units of *Taq* polymerase (Promega) in a total volume of 100 µl using a programme of 30 cycles at 1 min 95°C, 1 min 56°C and 2 min 72°C.

DNA from single calli and/or small leaf pieces was isolated according to Lassner *et al.* (1989). When using primer set NT1 and 1 the elongation time was extended to 5 min. The sequences of the primers used are: 1 5′-GAACTGACAGAACCGCAACG-3′; 10 5′-CATGCGATCATAGGCGTCTC-3′; MFA 5′-AGGGCCACAT-CAGCATCAAG-3′; MFB 5′-GTGGCCATTACCTCGTTAGC-3′; NT1 5′-CGCGGTGTCATCTATGTTAC-3′.

Southern analysis

Plant genomic DNA was isolated from young leaves harvested in the growth room and purified on a CsCl-gradient (Mettler *et al.*, 1987). Fifteen micrograms of DNA were digested with different enzymes and separated on a 0.7 or 1% agarose TBE gel. DNA was transferred to Hybond N+ membrane (Amersham) by capillary blotting using 0.4 M NaOH. The membrane was (pre-)hybridized according to the Hybond N^+ protocol. DNA probes labelled with [α-^{32}P]dCTP (specific activity: $0.5–1\times10^9$ d.p.m. μg^{-1} DNA) were obtained using a random primer labelling kit PRIME-IT (Stratagene) or a method described by Feinberg and Vogelstein (1983). Final washing was performed in 0.2×SSC, 0.1% SDS at 65°C. The blots were exposed to Fuji X-ray films or to Phosphorimager screens.

Acknowledgements

We would like to thank Roger Morton for critical reading of the manuscript. This work was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Technology Foundation (STW).

References

Adair, G.M., Nairn, R.S., Wilson, J.H., Seidman, M.M., Brotherman, K.A., MacKinnon, C. and Scheerer, J.B. (1989) Targeted homologous recombination at the endogenous adenine phosphoribosyltransferase locus in Chinese hamster cells. *Proc. Natl Acad. Sci. USA*, **86**, 4574–4578.

Aratani, Y., Okazaki, R. and Koyama H. (1992) End extension repair of introduced targeting vectors mediated by homologous recombination in mammalian cells. *Nucl. Acids Res.* **20**, 4795–4801.

Berinstein, N., Pennell, N., Ottaway, C.A. and Schulman, M.J. (1992) Gene replacement with one-sided homologous recombination. *Mol. Cell. Biol.* **12**, 360–367.

De Groot, M.J.A., Offringa, R., Does, M.P., Hooykaas, P.J.J. and van den Elzen, P.J.M. (1994) Non-recombinant background in gene targeting; Illegitimate recombination between a *hpt* gene and a defective 5′ deleted *nptII* gene can restore a Km^r phenotype. *Plant Mol. Biol.* **25**, 721–733.

Deng, C. and Capecchi, M.R. (1992) Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. *Mol. Cell. Biol.* **12**, 3365–3371.

Does, M.P., Dekker, B.M.M., de Groot, M.J.A. and Offringa, R. (1991) A quick method to estimate the T-DNA copy number in transgenic plants at an early stage after transformation, using inverse PCR. *Plant Mol. Biol.* **17**, 151–153.

Feinberg, A.P. and Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments for high specific activity. *Anal. Biochem.* **132**, 6–13, **137**, 266–267 (addendum 1984).

Fujiwara, T., Lessard, P.A. and Beachy, R.N. (1993) Inactivation of the nopaline synthase gene by double transformation: reactivation by segregation of the induced DNA. *Plant Cell Rep.* **12**, 133–138.

Gheysen, G., Villarroel, R. and van Montagu, M. (1991) Illegitimate recombination in plants: a model for T-DNA integration. *Genes Devel.* **5**, 287–297.

Halfter, U., Morris, P.-C. and Willmitzer, L. (1992) Gene targeting in *Arabidopsis thaliana*. *Mol. Gen. Genet.* **231**, 186–193.

Hasty, P., Rivera-Pérez, J. and Bradly, A. (1991) The length of homology required for gene targeting in embryonic stem cells. *Mol. Cell. Biol.* **11**, 5586–5591.

Horsch, R.B., Fry, J.E., Hoffmann, N.L., Eichholtz, D., Rogers, S.G. and Fraley, R.T. (1985) A simple and general method for transferring genes into plants. *Science*, **227**, 1229–1231.

Koncz, C., Németh, K., Rédei, G.P. and Schell J. (1994) Homology recognition during T-DNA integration into the plant genome. In *Homologous Recombination and Gene Silencing in Plants* (Paszkowski, J., ed.). Dordrecht: Kluwer Academic Publishers, pp. 167–189.

Lassner, M.W., Peterson, P. and Yoder, J.I. (1989) Simultaneous amplification of multiple DNA fragments by polymerase chain reaction in the analysis of transgenic plants and their progeny. *Plant Mol. Biol. Rep.* 7, 116–128.

Lee, K.Y., Lund, P., Lowe, K. and Dunsmuir, P. (1990) Homologous recombination in plant cells after *Agrobacterium*-mediated transformation. *Plant Cell*, **2**, 415–425.

Matsumoto, S., Ito, Y., Hosoi, T., Takahashi, Y. and Machida, Y. (1990) 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.

Matzke, M.A., Neuhuber, F. and Matzke, A.J.M. (1993) A variety of epistatic interactions can occur between partially homologous transgene loci brought together by sexual crossing. *Mol. Gen. Genet.* **236**, 379–386.

Mayerhofer, R., Konczkalman, Z., Nawrath, *et al.* (1991) T-DNA integration - a mode of illegitimate recombination in plants. *EMBO J.* **10**, 697–704.

Mettler, I.J. (1987) A simple and rapid method for minipreparation of DNA from tissue cultured plant cells. *Plant Mol. Biol. Rep.* **5**, 346–349.

Offringa, R. (1992) *Gene targeting in plants using the* Agrobacterium *vector system*. PhD Thesis Leiden University, The Netherlands.

Offringa, R., de Groot, M.J.A., Haagsman, H.J., Does, M.P., van den Elzen, P.J.M. and Hooykaas, P.J.J. (1990) Extra-chromosomal homologous recombination and gene targeting in plant cells after *Agrobacterium* mediated transformation. *EMBO J.* **9**, 3077–3084.

Offringa, R., van den Elzen, P.J.M. and Hooykaas, P.J.J. (1992a) Gene targeting in plants using the *Agrobacterium* vector system. *Transgenic Res.* **1,** 114–123.

Offringa, R., Franke-van Dijk, M.E.I., de Groot, M.J.A., van den Elzen, P.J.M. and Hooykaas, P.J.J. (1993) Nonreciprocal homologous recombination between *Agrobacterium* transferred DNA and a plant chromosomal locus. *Proc. Natl Acad. Sci. USA*, **90,** 7346–7350.

Ohl, S., Offringa, R., van den Elzen, P.J.M. and Hooykaas, P.J.J. (1994) Gene replacement in plants. In *Homologous Recombination and Gene Silencing in Plants* (Paszkowski, J., ed.). Dordrecht: Kluwer Academic Publishers, pp. 191–217.

Paszkowski, J., Baur, M., Bogucki, A. and Potrykus, I. (1988) Gene targeting in plants. *EMBO J.* **7,** 4021–4026.

Thomas, K.R., Deng, C. and Capecchi, M.R. (1992) High-fidelity gene targeting in embryonic stem cells by using sequence replacement vectors. *Mol. Cell. Biol.* **12,** 2919–2923.

Yenofsky, R.L., Fine, M. and Pellow, J.W. (1990) A mutant neomycin phosphotransferase II gene reduces the resistance of transformants to antibiotic selection pressure. *Proc. Natl Acad. Sci. USA*, **88,** 3435–3439.