Major Chromosomal Rearrangements Induced by T-DNA Transformation in Arabidopsis

Philippe Nacry, Christine Camilleri, Béatrice Courtial, Michel Caboche and David Bouchez

Laboratoire de Biologie Cellulaire, Institut National de la Recherche Agronomique, 78026 Versailles Cedex, France

Manuscript received December 15, 1997
Accepted for publication March 16, 1998

ABSTRACT

We show that major chromosomal rearrangements can occur upon T-DNA transformation of *Arabidopsis thaliana*. In the ACL4 line, two T-DNA insertion loci were found; one is a tandem T-DNA insert in a head-to-head orientation, and the other is a truncated insert with only the left part of the T-region. The four flanking DNA regions were isolated and located on the Arabidopsis chromosomes; for both inserts, one side of the T-DNA maps to chromosome *2*, whereas the other side maps to chromosome *3*. Both chromosome *3* flanking regions map to the same location, despite a 1.4-kb deletion at this point, whereas chromosome *2* flanking regions are located 40 cM apart on the bottom arm of chromosome *2*. These results strongly suggest a reciprocal translocation between chromosomes *2* and *3*, with the breakpoints located at the T-DNA insertion sites. The interchanged fragments roughly correspond to the 20-cM distal ends of both chromosomes. Moreover, a large inversion, spanning 40 cM on the genetic map, occurs on the bottom arm of chromosome *2*. This was confirmed by genetic analyses that demonstrated a strong reduction of recombination in the inverted region. Models for T-DNA integration and the consequences for T-DNA tagging are discussed in light of these results.

AGROBACTERIUM-MEDIATED T-DNA transformation (Zambryski *et al.* 1989) has been exploited for the construction of numerous plant transformation vectors in order to introduce new genes into plant cells (Lindsey and Topping 1996). A wide range of plant species are now routinely transformed using this system, including several monocots (reviewed in Park *et al.* 1996; Tinland 1996). In *Arabidopsis thaliana*, T-DNA has been used successfully as an insertional mutagen for gene tagging purposes, with over 40 genes isolated to date (reviewed in Lindsey and Topping 1996; Azpiroz-Leehan and Feldmann 1997).

The T-DNA transformation process itself has been extensively studied, especially the bacterial components involved in T-DNA mobilization and transfer, a phenomenon reminiscent of bacterial conjugation (reviewed in Lindsey and Topping 1996; Tinland 1996). Activation of the *vir* genes by plant compounds results in the synthesis of a linear, single-stranded copy of the T-DNA region associated with two types of vir proteins. The T-DNA is then translocated into the plant cell nucleus where it is stably integrated into the plant genome. Numerous T-DNA structures obtained after Agrobacterium-mediated transformation have been characterized in great detail, suggesting that integration sites are spread throughout the genome. In Arabidopsis, the mapped insertion loci are evenly distributed along the chromosomes (Azpiroz-Leehan and Feldmann 1997).

Corresponding author: David Bouchez, Laboratoire de Biologie Cellulaire, INRA, Route de Saint-Cyr, 78026 Versailles Cedex, France. E-mail: bouchez@versailles.inra.fr

In the plant genome, the right end of the T-DNA is frequently in the close vicinity of the 24-bp right border (RB) repeat, whereas the left end shows more variation, from a few to a few hundred nucleotides away from the 24-bp left border (LB) repeat (reviewed in Tinland 1996). Many lines of evidence suggest that T-DNA integration often induces base substitutions, insertions, and small (<500 bp) rearrangements (deletions and duplications) at the insertion site (Gheysen *et al.* 1991; Mayerhofer *et al.* 1991; Koncz *et al.* 1994). More significant chromosomal rearrangements have also been reported in tobacco by Ohba *et al.* (1995). In addition, Castle *et al.* (1993), studying 36 Arabidopsis embryo-defective mutants produced following seed T-DNA transformation, found indications for chromosomal translocations in nine of them. The authors concluded that chromosomal rearrangements could be a common feature of T-DNA transformed plants. Recently, Takano *et al.* (1997) described massive rearrangements (inversions and duplications) of genomic DNA at integration sites in rice.

Results obtained from these analyses have led several authors to propose different models for T-DNA integration. The first one is based on illegitimate recombination (Gheysen *et al.* 1991; Mayerhofer *et al.* 1991; Koncz *et al.* 1994; Tinland 1996). Recently, another model involving double strand break and repair has been proposed to account for the fact that T-DNA inserts were commonly found in tandem arrays at single or multiple loci (De Neve *et al.* 1997).

In this article, we describe a complex chromosomal structure induced by T-DNA insertion in an *Arabidopsis*

thaliana T-DNA mutagenized line. This line (ACL4) was selected during the visual screening of a T-DNA insertion population for morphological alterations. The T-DNA population was obtained by vacuum-infiltration transformation (Bechtold *et al.* 1993) using the Agrobacterium strain MP5-1 (Bouchez *et al.* 1993). Preliminary results showed that the mutation was apparently tagged by a T-DNA insert. Further characterization revealed complex chromosomal alterations associated with multiple T-DNA inserts: a reciprocal translocation between chromosomes *2* and *3*, a large inversion on the lower arm of chromosome *2*, and a short deletion on chromosome *3*. These rearrangements have been characterized at the molecular level and their consequences on genetic segregations, as well as different models that could lead to such complex chromosomal alterations are discussed.

MATERIALS AND METHODS

Arabidopsis lines and growing conditions: The ACL4 line derives from a T-DNA mutagenized population in ecotype Wassilevskija (WS), obtained by the vacuum-infiltration procedure (Bechtold *et al.* 1993), using the Agrobacterium strain MP5-1 carrying the transformation vector pGKB5 (Bouchez *et al.* 1993). Generations of T-DNA mutagenized plants are named as follows: T_0 = vacuum-treated plants; T_1 = Basta-resistant plants selected in T_0 progeny; T_2 = progeny of a T_1 selfed plant; T_3 = progeny of a pool of selfed T_2 plants.

For growth in the greenhouse, seeds were sown on soil and seedlings were transferred into individual pots 10 days after germination. Plants were grown on sterilized compost under 16 hr photoperiod, 10–15° night/20–25° day temperature.

For *in vitro* growth, seeds were surface sterilized and grown as described by Santoni *et al.* (1994). Kanamycin selection was performed on 100 μg/ml kanamycin-containing medium. Seeds were imbibed for at least 48 hr at 4° and transferred to a controlled-environment chamber (16 hr light/200 μmol photons m/sec, 70% relative humidity, 15° night temperature, 20° day temperature).

Oligonucleotides used for PCR: AKT1f (5′-ATGAGAGGAG GGGCTTTGTTATGCGG-3′); AKT1r (5′-CGAGGTAACCAA CAAAGAATGT-3′); T1Af (5′-GCGGTCTACTATCTTCATTT C-3′); T1Ar (5′-TTGGTTTCTGTAGGCTGAACT-3′); T1Bf (5′-TCCGACCATAGAGGATAAAATC-3′); T1Br (5′-ACGCT GCCTTGAGATAAACCA-3′); T2Af (5′-CATTTGATATTGT TAGTTGAAGTG-3′); T2Ar (5′-TTACATAGTAGAACAGAGA GGAT-3′); T2Bf (5′-CGACTCTGTTTCTGAATCTCTCC-3′); and T2Br (5′-TGTTTCTGCCGTATCCTCCTC-3′).

DNA extraction, PCR amplifications, CAPS analysis: Single mutant plant DNA was prepared as described by Edwards *et al.* (1991). The DNA pellet was dissolved in 50 μl H_2O and 2 μl was used for PCR amplification. Amplifications were carried out on a PTC100-96 thermal cycler (MJ Research, Inc., Watertown, MA) in 25 μl, containing 200 μmol of each dNTP, 2.5 mm of $MgCl_2$, 0.4 μmol of each primer and 0.2 units of Taq DNA polymerase. PCR conditions were as follows: 94°, 20 sec; 55°, 30 sec and 72°, 1 min; 37 cycles.

For CAPS analyses (Konieczny and Ausubel 1993), 5 μl of amplification product was digested by 2 units of restriction enzyme, and run on a 1.5% agarose gel. CAPS markers BGL1, GL1, GAPA, and m429 have been previously described by Konieczny and Ausubel (1993). AKT1f and AKT1r primers give a 1450-bp PCR product which is cleaved into two *Eco*RV fragments (1050 and 400 bp) only in ecotype Columbia. The cdc2Bf and cdc2Br (http://genome-www.stanford.edu/arabidopsis/maps/aboutcaps.html) give a 1400-bp PCR product in Columbia and a 900-bp PCR product in WS.

Southern analysis: DNA was isolated from four-week-old plants grown *in vitro* as previously described by Bouchez *et al.* (1996). Southern blot analysis was performed on 1 μg digested DNA, blotted onto Hybond N^+ membranes according to the manufacturer (Amersham, Buckinghamshire, UK). Probes were generated by the random primer method as described by Sambrook *et al.* (1989).

Genomic library: Two μg of plant DNA isolated from pooled mutant T_2 plants was digested to completion by *Eco*RI and then cloned into λ Zap II/*Eco*R1 Cloning Kit (Stratagene, La Jolla, CA) and packaged (Gigapack II Packaging extract; Stratagene) according to the manufacturer's instructions. DNA sequencing was performed using Taq DNA polymerase, dye-primers, and a ABI373A automated DNA sequencer. Experimental procedures were as recommended by the manufacturer (Applied Biosystems, Foster City, CA).

YAC library screening: Oligonucleotides were deduced from plant genomic sequences flanking T-DNA inserts. They were used for PCR screening of the CEPH/INRA/CNRS yeast artificial chromosome (CIC YAC) library as described by Creusot *et al.* (1995) on three-dimensional YAC pools. The length of the amplification products obtained with T1A, T1B, T2A, and T2B primers are respectively 689 bp, 136 bp, 208 bp, and 112 bp.

RESULTS

The ACL4 line was identified in a visual screen of the progeny of T_1 plants deriving from T_0 lines transformed with Agrobacterium strain MP5-1 carrying the transformation vector pGKB5 (Bouchez *et al.* 1993). The mutants, named *ton1* (Traas *et al.* 1995), are strongly compressed in the apical-basal axis and enlarged radially, resulting in dwarf, squat, and misshapen plantlets (Figure 1). Preliminary segregation analyses showed that the *ton1* phenotype appeared to be 100% linked to a T-DNA insert. On this basis, the ACL4 line was further studied to confirm this result.

Preliminary characterization of the ACL4 line: The structure of the T-DNA insert was determined by Southern analysis. DNA was extracted from pooled mutant T_2 plants, digested by seven restriction enzymes and hybridized to three different probes derived from the RB, LB, and central (KAN) parts of the T-region of pGKB5 (Figure 2). The hybridization patterns revealed two distinct T-DNA inserts (Figure 2). The first one (T-DNA1) is a tandem insert in inverted orientation with the left border of each T-DNA oriented toward plant genomic DNA. Both T-DNA copies appeared to be full-length inserts, with a total size of about 14 kb. The second insert (T-DNA2) is a truncated insertion (0.8 kb), consisting of the left part of the T-region, containing the 24-bp LB repeat and part of the Basta resistance gene.

Segregation analyses, performed on the pooled progenies of kanamycin resistant plants with wild-type phenotype (T_1 and T_2 generations) showed that in the ACL4 line, the mutation segregated with a 3:1 (wild-type:mutant) ratio, indicating a recessive, monogenic, nuclear

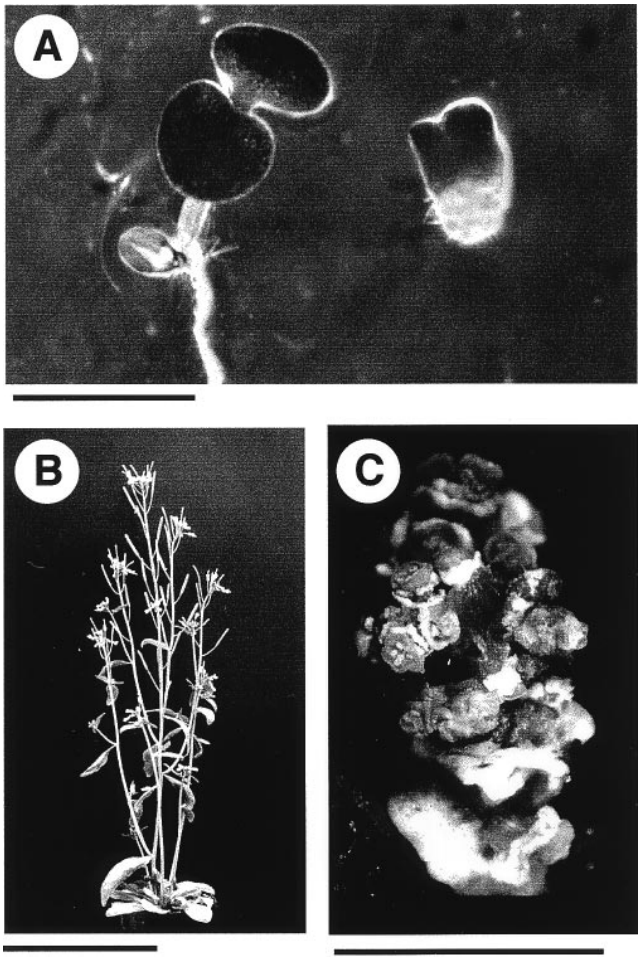

Figure 1.—Phenotype of *ton1* mutants. (A) Wild-type (left) and *ton1* (right) seedlings cultivated 7 days *in vitro*. (B) Wild-type plant cultivated 2 months in the greenhouse. (C) *ton1* mutant cultivated 2 months *in vitro*. Mutants are very short, thick, and misshapen, but all organs are in their correct relative positions, which can be readily recognized. All flower organs are present. Bars: (A) 1 cm; (B) 10 cm; (C) 1 cm.

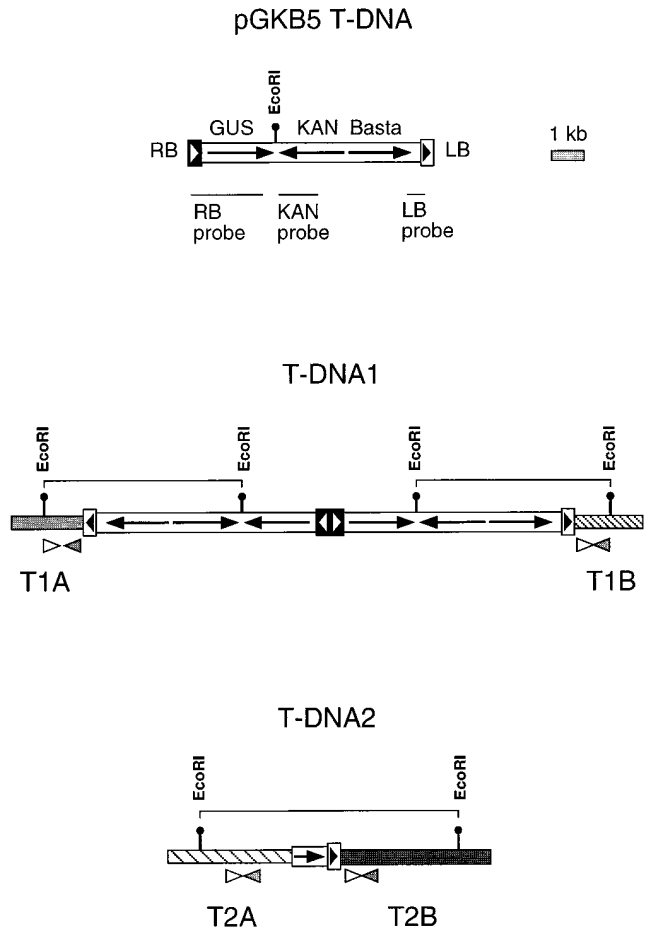

Figure 2.—Structure of pGKB5 T-DNA and of T-DNA insertions in the ACL4 line. The boxes containing arrowheads correspond to the 24-bp border sequences that serve as signals for T-DNA transfer. Arrows indicate coding sequences. GUS: β-glucuronidase; KAN: Tn5 kanamycin resistance under control of the nopaline synthase promoter; Basta: phosphinotrycin resistance. The extent of the RB, LB, and KAN probes used for Southern blots and library screening are indicated by bars. T-DNA1 is a head-to-head (RB-RB) tandem insert. T-DNA2 is a truncated insertion with only the left part of the pGKB5 T-DNA. Brackets indicate the different *Eco*RI fragments in the λ clones isolated from the mutant genomic library using the LB probe. T1A and T1B, T2A and T2B: genomic flanking sequences, respectively, for T-DNA1 and T-DNA2. Arrowheads indicate the location and the orientation of PCR primers used in this study (white: forward primer; grey: reverse primer). Solid circle: *Eco*RI sites.

mutation (Table 1). Moreover, in ACL4, the T-DNA segregated as a single insertion locus on the basis of the 3:1 kanamycin resistant:kanamycin sensitive ratio (Table 1).

Linkage between the T-DNA and the *ton1* mutation was tested by transferring 1427 mutants (882 T_2 and 545 T_3 plants) onto a kanamycin-containing medium: all of them were clearly kanamycin resistant. Moreover, PCR analysis of 245 mutant seedlings confirmed the presence of the tandem T-DNA insertion (T-DNA1). These results indicate a tight genetic linkage between the mutation and the T-DNA1.

In an attempt to separate the two T-DNA insertions, 148 T_3 plants with wild-type phenotype were individually grown in the greenhouse without any selection. Southern blot analysis was performed on *Eco*RI-digested DNA from individual plants. Hybridization with the LB probe (see Figure 2) revealed either no (57 plants) or three *Eco*RI junction fragments (91 plants), indicating tight genetic linkage of the two T-DNA inserts. Plants lacking T-DNA segregate 100% kanamycin sensitive seedlings, and mutants were never observed in their progeny.

Upon analysis of the segregation of kanamycin resistance (harbored by T-DNA1) and the *ton1* mutation in the progeny of 85 unpooled, individual T_2 and T_3 plants, we found significant deviations from the 3:1 (wild-type:mutant) ratio in most cases. Segregation data for 11 such lines are presented in Table 2. The lines could be classified into two different types. For most of the lines (type A), mutant frequency (from 22.5 to 40.9%) and kanamycin resistance frequency (from 75 to 80.8%) were higher than expected for monogenic Mendelian

TABLE 1

Genetic analysis of the *ton1* mutation and kanamycin resistance in the pooled progenies of T_2 and T_3 plants

No. of plants	*ton1*[a]	Wild type[a]	χ^2	Kanamycin resistant[a]	Kanamycin sensitive[a]	χ^2
2256	588 (26.1)	1668 (73.9)	1.36*	1727 (76.6)	529 (23.4)	2.89*

The χ^2 calculated values are based on an expected ratio of 3:1. * Value not significantly different at $P = 0.05$.

[a] Values in parentheses are percentages.

factors. About one third of the lines (type B) segregated approximately 50% kanamycin resistant and 50% sensitive plants, and less than 5% mutants (all kanamycin resistant). Such a segregation is maintained in subsequent generations when kanamycin resistant plants are selfed, and the mutant frequency is never higher than 5%. Plants exhibiting such segregation were always observed in the progeny of plants segregating kanamycin resistance and the mutation, even after three generations of backcrosses.

In pooled progenies (Table 3), plants deriving from type A and type B parent plants unexpectedly compensate for each other to give a segregation pattern that appears to be Mendelian for both kanamycin resistance and the *ton1* mutation (Table 1).

Molecular characterization of T-DNA insertions in ACL4: *Isolation and mapping of T-DNA flanking regions in ACL4:* With the aim of estimating the physical distance between the two T-DNA insertions, a genomic library of the mutant was constructed and screened using the LB probe (Figure 2). Nine positive clones were recovered, subcloned, and partially sequenced. Eight corresponded to the T-DNA1 flanking regions (three for T1A and five for T1B) and the ninth clone contained T-DNA 2 with genomic DNA on both sides (T2A and T2B) (Figure 2). Southern blot analysis and PCR amplifications confirmed that the isolated clones were colinear to the ACL4 genome (not shown). Synthetic oligonucleotides corresponding to the different genomic fragments were synthesized (Figure 2).

The four genomic sequences flanking the T-DNAs were mapped on the Arabidopsis YAC physical map. PCR primer pairs for each flanking sequence were used to screen the Arabidopsis CIC YAC library (Creusot *et al.* 1995). Both the T1A and T2B PCR primers identified six CIC YACs (9D7, 11B6, 8H7, 5D3, 12F10, and 7F7) located at the bottom of chromosome *3* (close to ve022; 73.5, cM-) (C. Camilleri and D. Bouchez, unpublished results). The T2A PCR primers identified four CIC YACs (9G1, 8B9, 7E2, and 5B4) located at the middle of chromosome *2* (between mi148 -36.3 cM- and m251 -39.5 cM-) (Zachgo *et al.* 1996). The T1B primers identified four CIC YACs (5E11, 6E11, 8H2, and 5C4) located at the bottom of chromosome *2* (close to SG5 -80.1 cM-) (Zachgo *et al.* 1996). RFLP analysis with a T1B probe on a set of 100 recombinant inbred (RI) lines (Lister and Dean 1993) confirmed the chromosomal location of this region (RFLP marker ve019 -81.4 cM-; Lister and Dean 1997).

Therefore, for both T-DNA insertions, flanking re-

TABLE 2

Genetic analysis of the *ton1* mutation and kanamycin resistance in the progeny of 11 (from 85 lines analyzed) individually selfed T_2 and T_3 heterozygous plants

Line no.	Type	No. of plants		*ton1* (%)	Wild type (%)	Kanamycin resistant (%)	Kanamycin sensitive (%)
		ton1	Wild type				
1	A	35	68	34.0	66.0	79.6	20.4
2	A	36	52	40.9	59.1	76.1	23.9
3	A	26	46	36.1	63.9	79.2	20.8
4	B	3	82	3.5	96.5	49.4	50.6
5	B	0	51	0.0	100.0	41.2	58.8
6	A	33	68	32.7	67.3	76.2	23.8
7	A	23	55	29.5	70.5	80.8	19.2
8	B	0	73	0.0	100.0	47.9	52.1
9	A	18	62	22.5	77.5	75.0	25.0
10	B	0	123	0.0	100.0	42.9	57.1
11	A	39	85	31.5	68.5	75.8	24.2

TABLE 3

Genetic analysis of the *ton1* mutation and kanamycin resistance in the progeny of individually selfed T_2 and T_3 heterozygous plants calculated on the progeny of 85 plants

Type of parent plants	No. of parent plants	No. of seedlings analyzed	*ton1* (%)	Wild type (%)		Kanamycin resistant (%)	Kanamycin sensitive (%)	
A	57	3605	1368 (37.9)	2237 (62.1)		2938 (81.5)	667 (18.5)	
B	28	1680	11 (0.7)	1669 (99.3)		783 (46.6)	897 (53.4)	
Total	85	5285	1379 (26.1)	3906 (73.9)	$\chi^2 = 2.8$*	3721 (70.4)	1564 (29.6)	$\chi^2 = 59.15$

Parent plants have been grouped on the basis of their segregation type (see Table 2). χ^2 calculated values are based on an expected ratio of 3:1. *χ^2 value not significantly different at $P = 0.05$.

gions map on chromosome *2* on one side (T1B and T2A), and on chromosome *3* on the other side (T1A and T2B) (Figure 2). Both chromosome *3* flanking regions map at the same location, whereas the chromosome *2* regions are about 40 cM apart on the lower arm. Southern analyses using T1A and T2B as probes on digests of WS DNA revealed common hybridizing restriction fragments (not shown), which confirmed their physical linkage in the ACL4 genome. These results strongly suggest that major chromosomal rearrangements (a reciprocal translocation between chromosome *2* and *3* on one hand, and a large deletion or inversion on chromosome *2* on the other hand) occurred in this line. Moreover, heterozygous ACL4 plants showed a significant reduction in pollen viability as tested by Alexander's staining method (Alexander 1969): 30–60% lethality, and also in ovule viability (~40% missing seeds in siliques), which is also a good indication for chromosomal defects.

A 1.4-kb deletion in chromosome 3: T1A and T2B T-DNA flanking regions were found on the same set of CIC YACs. Both T1A and T2B fragments were used as probes on restriction digests of these YAC clones, and they show the same hybridization pattern for the different restriction enzymes tested (not shown). Moreover, PCR amplifications were performed on wild-type genomic DNA using primers from each of the chromosome *3* flanking regions: T1Af and T2Br (Figure 2). A 2.2-kb fragment was amplified in wild-type DNA, instead of the 800 bp expected from the sequences of T1A and T2B, suggesting that a 1.4-kb region is deleted in the mutant DNA. Sequence comparison of the mutant and the wild-type genomic region confirmed the 1.4-kb deletion in ACL4.

Further genetic studies confirm the complex chromosome structure in ACL4: Taken together, the molecular results give strong indication for large chromosomal rearrangements, with breakpoints situated precisely at the T-DNA insertion sites (BP2b and BP3, Figure 3). As the T1B (T-DNA1) and the T2A (T-DNA2) flanking regions are located 40 cM apart, we tested several mutant plants for the presence of markers located in this interval to differentiate between a large deletion or, more likely, a large paracentric inversion on chromosome *2*. PCR analysis of 215 individual mutant plants was performed, using various primer combinations: two for each site, one specific for the wild-type sequence, one specific for the T-DNA insertion. This enabled us to distinguish all the different genomic structures at each T-DNA site, from wild-type homozygous to T-DNA homozygous. This analysis revealed that the T_2 mutant plants can be divided into two groups (Figure 4) according to their genomic composition: 70% of the mutant plants (group 1) are homozygous for both translocated chromosomes, whereas 30% (group 2) are homozygous for the translocated chromosome *3* and heterozygous for chromosome *2*.

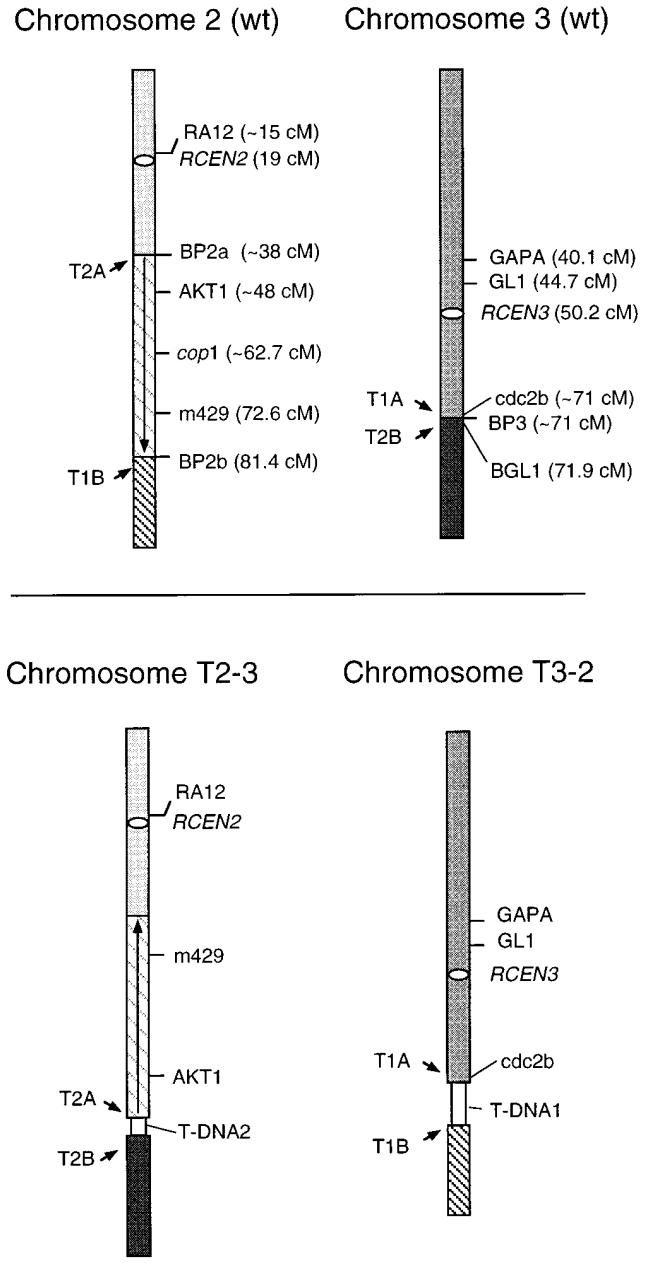

Figure 3.—Genetic maps of wild-type Arabidopsis chromosomes *2* and *3* and maps of rearranged chromosomes T2-3 and T3-2 in ACL4. All the genetic positions indicated refer to the reference RI map (Lister and Dean 1997). For the markers that are not present on the RI map, genetic locations were deduced from their position on the YAC physical map. RCEN2 and RCEN3 are centromeric repeat markers (Round *et al.* 1997). T1A and T1B, T2A and T2B: genomic flanking sequences, respectively, for T-DNA1 and T-DNA2. BP2a, BP2b, and BP3: chromosomal break points in ACL4. The white boxes indicate the T-DNA inserts. For clarity, the regions involved in the translocation and inversion are hatched differently (as in Figure 2), and the chromosome *2* inversion is indicated by an arrow.

As group 2 mutants carry one wild-type chromosome *2*, the distinction between a paracentric inversion and a large deletion on chromosome *2* was made from the analysis of isolated group 1 mutant plants (which carry two copies of a rearranged chromosome *2*). In case of a deletion, several chromosome *2* markers would simply be absent from these plants. We tested for the presence of several chromosome *2* PCR markers located in this region [ve061 (∼40 cM), AKT 1 (∼48 cM), m283 (60.4 cM), RB9 (∼62 cM), m429 (72.6 cM), ve065 (∼75 cM); Lister and Dean 1997; C. Camilleri and D. Bouchez, unpublished results] on DNA from 46 group 1 plants. All the markers tested in group 1 plants give an amplification product similar to the wild type. This demonstrates that the 40-cM region located between BP2a and BP2b on chromosome *2* is not deleted, but rather inverted (Figure 3).

To study the genetic behavior of such structures in Arabidopsis, we performed a linkage analysis on a population of F_2 mutant plants derived from a cross between ACL4 heterozygous plants (WS background) and Columbia wild-type plants. The resulting linkage data for different chromosome *2* and chromosome *3* markers are shown in Table 4. Three chromosome *2* genetic markers were tested, two in the inverted region (m429 and AKT1- map positions 72.6 and ∼48 cM-) and one in the nonrearranged region (RA12; ∼15 cM-) (Figure 3). The two markers located in the inverted region are linked to the *ton1* mutation and map at the same distance from the *ton1* mutation in our population, whereas they are about 25 cM apart on chromosome *2*. The RA12 marker is unlinked (Table 4).

Because we had previously found that one third of *ton1* mutants are heterozygous for the chromosome *2* rearrangement (group 2 plants), we analyzed the chromosomal structure of individual plants in our mapping population. Twenty-six group 2 mutants were eliminated from the mapping data set because these plants introduce a bias in the observed recombination frequency, due to the chromosomal heterozygous status. A new linkage analysis performed on the group 1 mutant population showed m429 and AKT1 still 100% linked to the *ton1* mutation, but RA12 at 18.1 cM (Table 4). These results show that recombination is suppressed in the inverted fragment and not affected in the region between the end of the inversion (BP2a) and the RA12 marker.

Four chromosome *3* markers were tested (Table 4), three of them are linked to the *ton1* mutation, with two (cdc2b and BGL1) 100% linked. The genetic distance (∼16 cM) between the *ton1* mutation and the third linked marker (GL1) is significantly reduced in our F_2 population compared to the genetic distance derived from the RI map (∼26 cM). In addition, no linkage is detected with the GAPA marker, which is only 5 cM north of GL1. These findings suggest that recombination is also modified in this chromosome *3* region in heterozygous plants. The chromosome *3* genetic distances recalculated on the group 1 population are unchanged.

In addition, the *cop1-6* (Deng *et al.* 1991; chromosome *2*, 62.7 cM) mutation was used as a phenotypic marker.

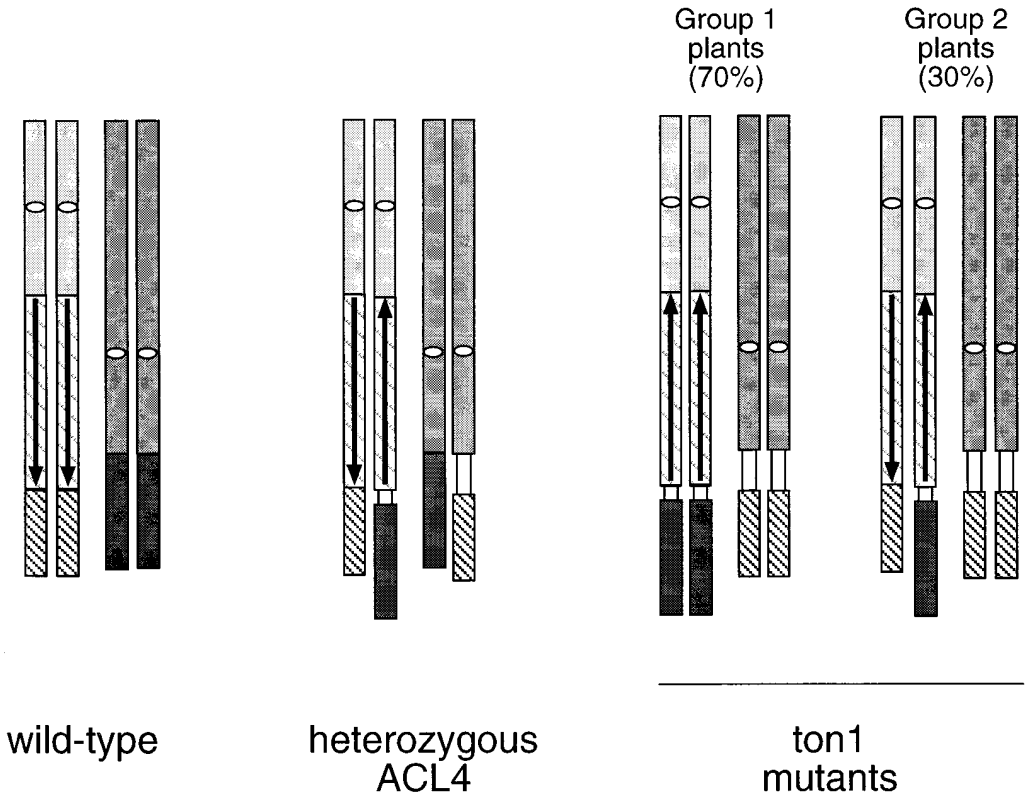

Figure 4.—Chromosomal structure in wild-type, heterozygous, and *ton1* mutant plants. The ACL4 heterozygous plants contain for both chromosomes *2* and *3* one wild-type and one rearranged chromosome (Chromosomes *2*; T2-3 and *3*; T3-2). The *ton1* mutants isolated in the progeny of selfed heterozygous plants can be divided into two groups: group 1 plants (70%) are homozygous for both translocated chromosomes (chromosomes T2-3; T2-3 and T3-2; T3-2) whereas group 2 plants (30%) are homozygous for the translocated chromosome *3* and heterozygous for chromosome *2* rearrangement (chromosomes *2*; T2-3 and T3-2; T3-2). Motives used to identify the different chromosomal regions are the same as in Figures 1 and 2.

In the F_2 progeny of a cross between *ton1* heterozygous plants and *cop1-6* mutants, only two of the 362 *cop1-6* mutant seedlings were kanamycin resistant. This further substantiates previous results showing that recombination frequencies are strongly altered on the distal arm of chromosome *2*.

DISCUSSION

The characterization of numerous T-DNA insertion sites in chromosomes of several plant species has been reported. Most insertions correspond to simple, unique inserts where both T-region ends are present, with only short deletions removing just a few nucleotides in plant DNA at the integration site (Gheysen *et al.* 1991; Mayerhofer *et al.* 1991; Koncz *et al.* 1994). These authors also noted short stretches of similarity between the T-DNA ends and the preintegration site, and suggested that breaks in chromosomal DNA were required for insertion of the T-DNA. Therefore, it has been proposed that T-DNA integration into plant chromosomes may involve a process of illegitimate recombination (reviewed in Gheysen *et al.* 1991; Mayerhofer *et al.* 1991; Koncz *et al.* 1994; Tinland 1996). More complex cases of T-DNA integration, such as deletions or duplications of T-DNA extremities or of genomic sequences (Gheysen *et al.* 1991; Castle *et al.* 1993; Ohba *et al.* 1995), as well as the presence of sequences of unknown origin between the T-DNA ends and genomic DNA (Gheysen *et al.* 1991) have also been described.

In ACL4, the first insertion (T-DNA1) corresponds to a head to head (RB-RB) tandem T-DNA. Such multiple T-DNA inserts, in various configurations, appear to be quite frequent (Castle *et al.* 1993; Koncz *et al.* 1994; Cluster *et al.* 1996; De Neve *et al.* 1997), and De Neve *et al.* (1997) found a preference for RB-RB or RB-LB associations. On the contrary, the second T-DNA insertion (T-DNA2) is a truncated insert, with only the left part of the T-region inserted. This kind of structure is poorly documented, probably because most studies are performed on plants selected on the basis of antibiotic or herbicide resistance (*i.e.*, a selection for at least one complete or nearly complete T-DNA insert). Castle *et al.* (1993) found that about one fourth of the T-DNA embryonic mutants they analyzed contained truncated T-DNAs, most of them corresponding to the left part of the T-region. These results support the hypothesis that T-DNA integration should be initiated on the LB side of the T-DNA, as proposed by Tinland (1996). In addition, in ACL4, no T-DNA fragment was detected by hybridization at break point BP2a. However, abortive T-DNA insertions may be involved in the formation of such chromosome breakages. This could explain the deletions, additions, and base substitutions observed in untagged mutants isolated from T-DNA mutagenized populations (Negruk *et al.* 1996), in which up to 65% of the observed mutations do not appear to be caused by a complete T-DNA insert (Azpiroz-Leehan and Feldmann 1997).

In the ACL4 line, T-DNA insertion induced large chromosomal rearrangements: a reciprocal translocation (interchange of the 20-cM distal ends of chromo-

TABLE 4

Genetic distances between the *ton1* mutation and chromosome *2* and *3* markers

Marker	Chromosome (position in cM)	Recombination percent[a]	Genetic distance to *ton1* (cM)	No. of plants scored
RA12	*2* (∼15)	45.8 ± 5.8	Unlinked	72 group 1 + 2
		17.4 ± 5.5	18.1 ± 6.3	46 group 1
AKT1	*2* (∼48)	18.1 ± 4.5	18.9 ± 5.2	72 group 1 + 2
		0.0	0.0	46 group 1
m429	*2* (72.6)	18.1 ± 4.5	18.9 ± 5.2	72 group 1 + 2
		0.0	0.0	46 group 1
GAPA	*3* (40.1)	47.2 ± 5.8	Unlinked	72 group 1 + 2
		47.8 ± 7.4	Unlinked	46 group 1
GL1	*3* (44.7)	17.9 ± 4.3	16.5 ± 4.8	72 group 1 + 2
		15.2 ± 5.2	15.7 ± 5.8	46 group 1
cdc2b	*3* (∼71)	0.0	0.0	72 group 1 + 2
		0.0	0.0	46 group 1
BGL1	*3* (71.9)	0.0	0.0	72 group 1 + 2
		0.0	0.0	46 group 1

An ACL4 heterozygous plant (WS background) and a Columbia wild-type plant were crossed. Seventy-two seedlings homozygous for the *ton1* mutation (group 1 + 2) were selected in the F_2 population deriving from this cross. The chromosomal structure of these mutants was analyzed by PCR and 26 of them were found to carry one wild-type chromosome *2* (group 2 plants). CAPS markers were tested on all F_2 plants. To estimate the genetic distances in this cross between the *ton1* mutation and several markers from chromosome *2*, and to compare these distances to those derived from the reference RI map (Lister and Dean 1997), we had to remove group 2 plants from the analysis. The genetic distances and standard deviation were calculated using the Kosambi mapping function (Koornneef and Stam 1992).

[a] Values are ±SD.

somes *2* and *3*), and a 40-cM inversion on chromosome *2*. BP2b and BP3 break points are precisely located at the T-DNA insertion sites. Few reports of such complex chromosomal rearrangements exist, possibly because in Arabidopsis, most chromosomal rearrangements leading to partial aneuploidy/polyploidy are expected to be lethal, due to the small size of the Arabidopsis genome which has small intergenic and noncoding regions and low gene redundancy.

Genetic results from Castle *et al.* (1993) suggested that about 20% of the T-DNA mutants examined (selected for embryo lethality) contained translocations or inversions, and Feldmann *et al.* (1997) also suggested the occurrence of translocations in T-DNA lines to account for exceptional segregations of the T-DNA markers. Takano *et al.* (1997) described massive rearrangements of genomic DNA at the foreign DNA integration sites in rice. Recently, Ray *et al.* (1997) reported on the characterization of a reciprocal chromosomal translocation in an Arabidopsis line isolated from a T-DNA mutagenized population.

In view of the current models for T-DNA integration into the plant genome (Gheysen *et al.* 1991; Mayerhofer *et al.* 1991; Koncz *et al.* 1994; Tinland 1996), several hypotheses can be made concerning the origin of the chromosomal structure observed in the ACL4 line. However, experimental confirmation of any of these hypotheses would be extremely difficult. The reciprocal translocation between chromosomes *2* and *3* may have arisen from homologous recombination between independent T-DNA inserts on these chromosomes. Recombination events between copies of transposable elements are known to induce chromosomal rearrangements (reviewed in Robbins *et al.* 1989), and homologous recombination between a T-DNA insert and either another T-DNA, or plant genomic DNA, has previously been reported (De Neve *et al.* 1997). Nevertheless, the occurrence of an additional large inversion on chromosome *2* would require a second independent recombination event which would be highly unlikely. In addition, no T-DNA remnant is observed at break point BP2a, and small deletions occur at break points. These findings do not favor a recombination mechanism and therefore events involving break and repair phenomena are more likely to be responsible. In this case, the insertion of T-DNAs 1 and 2 could have induced major rearrangements, due to defective repair mechanisms. The involvement of double strand breaks in T-DNA integration, already reported by De Neve *et al.* (1997), could account for the deletions observed at the plant/T-DNA junctions. These types of chromosomal rearrangements are known to induce particular structures at meiosis, which strongly affects gamete type and frequency. Reciprocal translocations have been extensively analyzed in rye and maize (reviewed in Sybenga 1972). In these plants, it has been shown that all homologous segments of a heterozygous reciprocal translocation can pair at meiosis and form a cross-shaped configuration of four chromosomes.

In the progeny of ACL4 heterozygous plants, we

found significant deviation from Mendelian segregation in the progeny of individual heterozygous plants (Table 2). Such defects have been observed in the progeny of plants carrying chromosomal translocations. Theoretically, the different chromosomal configurations should occur with equal frequencies, but in fact this depends on the presence and number of chiasmata, centromere activity, the size of the interchanged segments and chromosome flexibility (Sybenga 1972). Inversions are also known to induce gamete lethality (Redei and Koncz 1992). Unbalanced gametes resulting from adjacent chromosome segregation are generally lethal and induce semisterility in plants (Sybenga 1972; Koornneef 1994). In the ACL4 line, all the unbalanced gametes are apparently lost except one gametic genotype (chromosomes *2*, T3-2; Figures 2 and 3) resulting from an adjacent segregation.

Transmission defects can account for the biases in kanamycin resistance and mutant frequencies, but not for the near 1:1 Kan^R:/Kan^S segregation observed in the progeny of type B plants. These plants have both T-DNAs but mutants are rarely observed in their progeny. Crosses with wild-type plants demonstrate significant defects in kanamycin resistance transmission (12% by pollen and 39% by ovule, instead of 50%). In addition, the selfed progeny of these kanamycin resistant plants segregate approximately 50% Kan^R and 50% Kan^S plants. The heritability of these genetic segregations indicates the presence of additional chromosomal alterations that remain to be characterized.

We also found that the ACL4 chromosomal rearrangements induce modifications in crossing-over frequency. Genetic distances are significantly reduced in the vicinity of BP3 in the nontranslocated chromosomal fragment (GL1 marker). This may be due to the relatively short distance between the break point and the centromere (crossing-over frequency is usually reduced near the centromere). In addition, crossing-over in the interstitial fragment (between the centromere and the translocation point) frequently leads to unbalanced gametes that are lost (Sybenga 1972). On the contrary, we have observed a low level of recombination in the inverted chromosome *2* fragment (*cop1-6* marker). The large size of the inverted fragment (40 cM) may allow the formation of a loop allowing pairing with the nonrearranged homologous chromosome (Redei and Koncz 1992).

We now have convincing evidence that T-DNA integration can provoke profound rearrangements in plant genomes, both at the chromosomal level and at the gene level. The prevalence of large chromosomal alterations in T-DNA transformants is difficult to assess, but the actual frequency of such events could be significant and needs further examination. Whether this observation can be generalized to other transformation systems, such as direct transformation procedures, is unknown at present. We can expect that Agrobacterium-mediated transformation, as a highly specialized system, has evolved to recruit host functions involved in DNA repair mechanisms. The analysis of chromosomal structure in the ACL4 line provides further evidence that break and repair mechanisms are involved in the T-DNA integration process. Multiple T-DNA insertions, either successful or abortive, are then likely to generate large chromosomal rearrangements, such as those observed in ACL4.

From an evolutionary point of view, our results leave open the possibility that T-DNA induced chromosomal rearrangements may in some cases play a role in genome evolution and speciation. T-DNA is a natural agent of mutagenesis in plants (although probably not frequent in Arabidopsis), and a T-DNA remnant has been detected in the evolution of the genus Nicotiana (Furner *et al.* 1986). Natural T-DNAs carry oncogenes and opine synthesis genes that are most probably incompatible with normal growth and development. However, it would be interesting to test whether portions of Agrobacterium T-DNA, as described here, can be detected in natural populations.

T-DNA insertional mutagenesis is broadly used to generate mutants useful for gene cloning and functional analysis in plants. The occurrence of large chromosomal rearrangements in T-DNA lines, which can involve multiple loci not physically linked to one another, may strongly hamper the molecular characterization of putatively tagged mutations. In addition, at least some of the many untagged mutations observed in T-DNA mutagenized populations could be due to unprecise repairs associated with abortive or truncated T-DNA integration. Evidence of distortions in segregation ratios, and of semisterility in pollen or ovule development, can give good indications of chromosomal rearrangements. In particular, as shown here, genetic analyses of pooled progenies can be misleading. However, as far as Arabidopsis is concerned, the availability of powerful tools for genetic and physical mapping, and in the near future of the complete genome sequence, provides invaluable help in the identification of the regions involved in any mutant phenotype.

We thank J. Goujaud and J. Talbotec for taking care of the transgenic plants and H. McKhann, P. Vittorioso and H. Vaucheret for critical reading of the manuscript. This work was supported by the French Ministère de la Recherche et de l'Enseignement Supérieur grant no. 94245.

LITERATURE CITED

Alexander, M. P., 1969 Differential staining of aborted and non aborted pollen. Stain Technol. **44:** 117–122.

Azpiroz-Leehan, R., and K. A. Feldmann, 1997 T-DNA insertion mutagenesis in Arabidopsis: going back and forth. Trends Genet. **13:** 152–156.

Bechtold, N., J. Ellis and G. Pelletier, 1993 *In planta* Agrobacterium mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. C. R. Acad. Sci. Paris, Life Sci. **316:** 1194–1199.

Bouchez, D., C. Camillieri and M. Caboche, 1993 A binary vector based on Basta resistance for *in planta* transformation of *Arabidopsis thaliana.* C. R. Acad. Sci., Life Sci. **316:** 1188–1193.

Bouchez, D., P. Vittorioso, B. Courtial and C. Camillieri, 1996

Kanamycin rescue: a simple technique for the recovery of T-DNA flanking sequences. Plant Mol. Biol. **14:** 115–123.

Castle, L. A., D. Errampalli, T. L. Atherton, L. H. Franzmann, E. S. Yoon *et al.*, 1993 Genetic and molecular characterisation of embryonic mutants identified following seed transformation in Arabidopsis. Mol. Gen. Genet. **241:** 504–514.

Cluster, P. D., M. Odell, M. Metzlaff and R. B. Flavell, 1996 Details of T-DNA structural organization from a transgenic Petunia population exhibiting co-suppression. Plant Mol. Biol. **32:** 1197–1203.

Creusot, F., E. Fouilloux, M. Dron, J. Lafleuriel, G. Picard *et al.*, 1995 The CIC library: a large insert YAC library for genome mapping in *Arabidopsis thaliana*. Plant J. **8:** 763–770.

De Neve, M., S. Debuck, A. Jacobs, M. Van Montagu and A. Depicker, 1997 T-DNA integration patterns in co-transformed plant cells suggest that T-DNA repeats originate from co-integration of separate T-DNAs. Plant J. **11:** 15–29.

Deng, X. W., T. Caspart and P. H. Quail, 1991 *cop1*: a regulatory locus involved in light-controlled development and gene expression in Arabidopsis. Genes Dev. **5:** 1172–1182.

Edwards, K., C. Johnstone and C. Thompson, 1991 A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Res. **19:** 1349.

Feldmann, K. A., D. A. Coury and M. L. Christianson, 1997 Exceptional segregation of a selectable marker (Kan^R) in Arabidopsis identifies genes important for gametophytic growth and development. Genetics **147:** 1411–1422.

Furner, I. J., G. A. Huffman, R. M. Amasino, D. J. Garfinkel, M. P. Gordon *et al.*, 1986 An Agrobacterium transformation in the evolution of the genus Nicotiana. Nature **319**: 422–427.

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

Koncz, C., K. Németh, G. P. Redei and J. Schell, 1994 Homology recognition during T-DNA integration into the plant genome, pp. 167–189 in *Homologous Recombination and Gene Silencing in Plants*, edited by J. Paszkowski. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Konieczny, A., and F. M. Ausubel, 1993 A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. Plant J. **4:** 403–410.

Koornneef, M., 1994 Arabidopsis genetics, pp. 89–120 in *Arabidopsis*, edited by E. M. Meyerowitz and C. R. Sommerville. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Koornneef, M., and P. Stam, 1992 Genetic analysis, pp. 83–99 in *Methods in Arabidopsis Research*, edited by C. Koncz, N.-H. Chua and J. Schell. World Scientific, Singapore.

Lindsey, K., and J. F. Topping, 1996 T-DNA-mediated insertional mutagenesis, pp. 275–300 in *Plant Gene Isolation: Principles and Practice*, edited by G. D. Foster and D. Twell. John Wiley & Sons, New York.

Lister, C., and C. Dean, 1993 Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. Plant J. **4:** 745–750.

Lister, C., and C. Dean, 1997 The latest RI map using Lister and Dean RI lines. Weeds World (http://nasc.nott.ac.uk:8300/Vol4i/) **4(i):** 1–10.

Mayerhofer, R., Z. Koncz-Kalman, C. Nawrath, G. Bakkeren, A. Crameri *et al.*, 1991 T-DNA integration: a mode of illegitimate recombination in plants. EMBO J. **10:** 697–704.

Negruk, V., G. Eisner and B. Lemieux, 1996 Addition-deletion mutations in transgenic *Arabidopsis thaliana* generated by the seed co-cultivation method. Genome **39:** 1117–1122.

Ohba, T., Y. Yoshioka, C. Machida and Y. Machida, 1995 DNA rearrangements associated with the integration of T-DNA in tobacco: an example for multiple duplications of DNA around the integration target. Plant J. **7:** 157–164.

Park, S. H., S. R. M. Pinson and R. H. Smith, 1996 T-DNA integration into genomic DNA of rice following Agrobacterium inoculation of isolated shoot apices. Plant Mol. Biol. **32:** 1135–1148.

Ray, S., S. S. Park and A. Ray, 1997 Pollen tube guidance by the female gametophyte. Development **124:** 2489–2498.

Redei, G. P., and C. Koncz, 1992 Classical mutagenesis, pp. 16–82 in *Methods in Arabidopsis Research*, edited by C. Koncz, N.-H. Chua and J. Schell. World Scientific, Singapore.

Robbins, T. P., R. Carpenter and E. S. Coen, 1989 A chromosome rearrangement suggests that donor and recipient sites are associated during Tam3 transposition in *Antirrhinum majus*. EMBO J. **8:** 5–13.

Round, E. K., S. K. Flowers and E. J. Richards, 1997 *Arabidopsis thaliana* centromere regions: genetic map positions and repetitive DNA structure. Genome Res. **7:** 1045–1053.

Sambrook, J., E. F. Fritsch and T. Maniatis, 1989 *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Santoni, V., C. Bellini and M. Caboche, 1994 Use of two-dimensional protein-pattern analysis for three characterizations of *Arabidopsis thaliana* mutants. Planta **192:** 557–566.

Sybenga, J., 1972 *General Cytogenetics*. North Holland/American Elsevier, New York.

Takano, M., H. Egawa, J. E. Ikeda and K. Wakasa, 1997 The structures of integration sites in transgenic rice. Plant J. **11:** 353–361.

Tinland, B., 1996 The integration of T-DNA into plant genomes. Trends Plant Sci. **1:** 178–183.

Traas, J., C. Bellini, P. Nacry, J. Kronenberger, D. Bouchez *et al.*, 1995 Normal differentiation patterns in plants lacking microtubular preprophase bands. Nature **375:** 676–677.

Zachgo, E. A., M. L. Wang, J. Dewdney, D. Bouchez, C. Camilleri *et al.*, 1996 A physical map of chromosome *2* of *Arabidopsis thaliana*. Genome Res **6:** 19–25.

Zambryski, P., J. Tempe and J. Schell, 1989 Transfer and function of T-DNA genes from Agrobacterium Ti and Ri plasmids in plants. Cell **56:** 193–201.

Communicating editor: D. Preuss