The Plant Journal (1998) **13**(5), 707–716

TECHNICAL ADVANCE

Gene identification with sequenced T-DNA tags generated by transformation of *Arabidopsis* cell suspension

Jaideep Mathur[1], László Szabados[2], Sabine Schaefer[1], Britta Grunenberg[1], Andrea Lossow[1], Esther Jonas-Straube[1], Jeff Schell[1], Csaba Koncz[1,2] and Zsuzsanna Koncz-Kálmán[1,*]

[1]Max-Planck Institut für Züchtungsforschung, D-50829 Köln, Carl-von-Linné-Weg 10, Germany, and
[2]Institute of Plant Biology, Biological Research Center of Hungarian Academy of Sciences, H-6701 Szeged, PO Box 521, Temesvári krt. 62, Hungary

Summary

A protocol for establishment and high-frequency *Agrobacterium*-mediated transformation of morphogenic *Arabidopsis* cell suspensions was developed to facilitate saturation mutagenesis and identification of plant genes by sequenced T-DNA tags. Thirty-two self-circularized T-DNA tagged chromosomal loci were isolated from 21 transgenic plants by plasmid rescue and long-range inverse polymerase chain reaction (LR-iPCR). By bidirectional sequencing of the ends of T-DNA-linked plant DNA segments, nine T-DNA inserts were thus localized in genes coding for the *Arabidopsis* ASK1 kinase, cyclin 3b, J-domain protein, farnesyl diphosphate synthase, ORF02, an unknown EST, and homologues of a copper amine oxidase, a peripheral Golgi protein and a maize pollen-specific transcript. In addition, 16 genes were identified in the vicinity of sequenced T-DNA tags illustrating the efficiency of genome analysis by insertional mutagenesis.

Introduction

Following large-scale sequencing of expressed sequence tags (ESTs) and construction of physical maps, genome projects in *Arabidopsis* and other plants are advancing towards sequencing of whole chromosomes in conjunction with functional analysis of the sequenced genes (Cooke *et al.*, 1996; Höfte *et al.*, 1993; Huang and Miao, 1997; Newman *et al.*, 1994; Sasaki *et al.*, 1994; Schmidt *et al.*,1995; Zachgo *et al.*, 1996). As gene replacement is not yet routine in higher plants, the T-DNA of *Agrobacterium* and transposons are used for saturating the genome with insertional mutations in order to establish correlations between sequence data, mutant phenotypes and gene functions by reverse genetics (for reviews see Feldmann *et al.*, 1994; Long *et al.*, 1993; Miao and Lam, 1995; Osborn *et al.*, 1991; Souer *et al.*, 1995). Application of reporter gene fusion, enhancer trap, suicide gene and activator tagging technologies facilitate both targeted tagging of single genes and mutagenesis of particular gene classes characterized by well-defined regulatory functions and/or expression patterns (Franzmann *et al.*, 1995; Hayashi *et al.*, 1992; Kertbundit *et al.*, 1991; Knapp *et al.*, 1994; Koncz *et al.*, 1989, 1994; Sunderesan *et al.*, 1995; Topping *et al.*, 1991). In addition, PCR screening techniques adopted from *Drosophila* and *Caenorhabditis* genome research are used in the sequence-based identification of transposon and T-DNA insertions in known genes (Ballinger and Benzer, 1989; Kaiser and Goodwin, 1990; Koes *et al.*, 1995; Krysan *et al.*, 1996; McKinney *et al.*, 1995; Rushforth *et al.*, 1993; Souer *et al.*, 1995; Zwaal *et al.*, 1993).

Thus far, partial sequencing of 30 000 cDNAs in *Arabidopsis* has identified about 10 000 genes, corresponding to 40–65% of the total gene number, as calculated on the basis of genome size, mRNA complexity and mutation rates (Koncz and Rédei, 1994; Meyerowitz, 1994; Miklos and Rubin, 1996). Although the construction of equalized cDNA libraries may further increase this number (Kohchi *et al.*, 1995), the identification of all *Arabidopsis* genes now requires a link between genome sequencing and reverse genetics. Sequencing of plant DNA junctions of transposon and T-DNA tags may in fact provide such a link, assuming that saturation mutagenesis can be achieved with these insertional mutagens. Currently, one- and two-element systems based on the maize transposons *Ac/Ds* and *En/I/dSpm* (Bancroft *et al.*, 1993; Cardon *et al.*, 1993) are used to build up mutagenized populations, and already two large collections of T-DNA tagged lines, consisting of about 55 000 plants, are available in *Arabidopsis* (Bechtold *et al.*, 1993; Forsthoefel *et al.*, 1992). Genome saturation with T-DNA inserts is a realistic goal, because within a pool of 75 000–100 000 transgenic plants, a T-DNA tag is expected to occur in each *Arabidopsis* gene at 95–97.5% probability when the average copy number of inserts is between 1 and 3 (Bechtold *et al.*, 1993; Feldmann *et al.*, 1994).

Saturation mutagenesis with T-DNA inserts can be reproducibly achieved by *Agrobacterium* transformation of *Arabidopsis* cell suspensions as described below. In addition to promoting the development of targeted gene tagging

Received 8 July 1997; revised 30 October 1997; accepted 7 November 1997.
*For correspondence (fax +49 221 5062 213; e-mail koncz@mpiz-koeln.mpg.de).

approaches (e.g. activator T-DNA tagging; Kakimoto, 1996), this technique is applicable for mass regeneration of transgenic plants to generate mutant pools for identification of T-DNA-tagged *Arabidopsis* chromosomal loci by genome sequencing. Plasmid rescue and long-range inverse PCR offer two simple methods for isolation by self-circularization and bi-directional sequencing of T-DNA-tagged chromosomal DNA segments, facilitating the analysis of the *Arabidopsis* genome by reverse genetics.

Results and discussion

For *Agrobacterium*-mediated transformation of *Arabidopsis*, a wide range of tissue culture and *in planta* transformation methods are available. From *Agrobacterium*-infected leaf discs, cotyledon, stem or root explants, transgenic calli can be obtained and regenerated to plants with varying efficiencies (for review see Koncz *et al.*, 1994; Morris and Altmann, 1994). Because *Agrobacterium* is capable of systemically transforming diverse cell types when infiltrated into plants (e.g. as was demonstrated using a T-DNA with an intron-containing *uidA* reporter gene by Vancanneyt *et al.*, 1990), the need for plant regeneration from tissue culture can be overcome by *in planta* transformation. Because the yield of current transformation methods is generally a few thousand transgenic plants per experiment, saturation of the *Arabidopsis* genome with T-DNA tags is a cumulative process leading to the construction of a single representative mutant collection (for review see Bechtold *et al.*, 1993; Feldmann *et al.*, 1994). However, when mutant selection techniques are applicable at the cellular level using, e.g. activator mutagenesis (Hayashi *et al.*, 1992), saturation T-DNA tagging can be achieved in a single experiment by co-cultivation of plant protoplasts with *Agrobacterium* (Horsch *et al.*, 1987). In *Arabidopsis*, this technology is not yet efficient when applying the available leaf mesophyll- and root-derived protoplast systems (Damm *et al.*, 1989; Mathur *et al.*, 1995). However, similarly to tobacco and rice (An, 1985; Hiei *et al.*, 1994), *Agrobacterium* infection of morphogenic *Arabidopsis* cell suspensions yields very high transformation rates.

Initiation and maintenance of Arabidopsis *cell suspensions*

A protocol for the establishment of cell suspensions from *Arabidopsis* ecotypes Col-1, Col-5, RLD1 and WS2 is described in Experimental procedures. Briefly, root cultures were initiated after germination of sterilized seeds for a week by growing 15–20 seedlings in Erlenmeyer flasks containing basal medium (BM: regular MS medium (Murashige and Skoog, 1962) with B5 vitamins, Gamborg *et al.*, 1968) for 15–20 days. Roots were excised from seedlings, dissected and cultured in CM (callus medium, BM supplemented with auxins and cytokinin to promote cell division) in the dark for 15–20 days. Cell proliferation was initiated at the sites of lateral root initials, and more extensively, in cortical tissues of roots (Figure 1a). When seeds were directly germinated in CM, cell proliferation was also inducible in the hypocotyls (Figure 1b). Proliferating cell clumps released in the medium were separated from explants (which were further cultured to raise more material) using a sieve with a pore size of 850 μm, concentrated, and propagated for 5 weeks by weekly sub-culturing. Finally, the cultures were filtrated through a sieve of 250 μm pore size to obtain quickly cycling fine cell suspensions (Figure 1c).

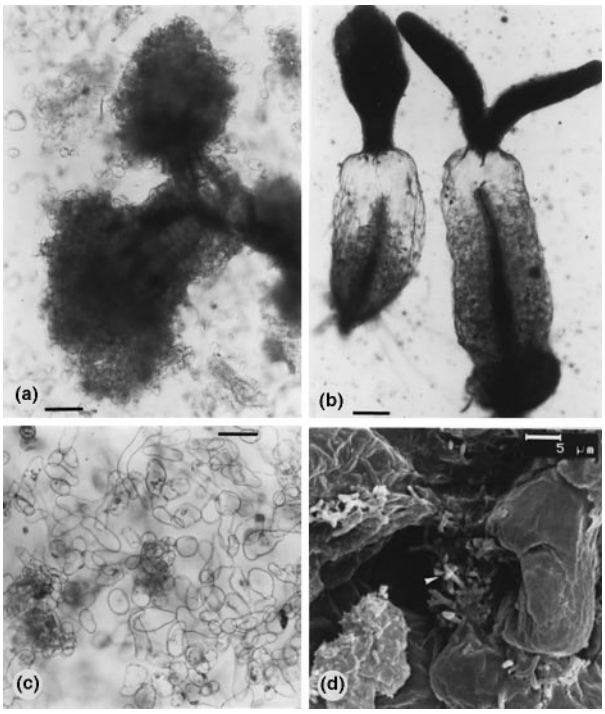

Figure 1. Establishment of cell suspension from root explants and seedlings of *Arabidopsis thaliana* for co-cultivation with *Agrobacterium*.
(a) Proliferating cells on root explants after 7 days of culture in liquid CM.
(b) Cell proliferation in the hypocotyl of seedlings in CM.
(c) Cell suspension obtained after four sub-cultures (35 days) and filtration through a sieve of 250 μm.
(d) Scanning electromicrograph showing *Agrobacterium* attachment to the surface of cells in suspension 5 h after co-cultivation.
Bars: (a) 0.25 mm; (b) 0.05 mm; (c) 25 μm; (d) 5 μm.

Combined application of 2,4-D (2,4-dichlorophenoxyacetic acid) and IAA (indole-3-acetic acid) was found to promote faster cell division in most *Arabidopsis* ecotypes than 2,4-D alone. The size of cell clumps obtained with 2,4-D and IAA ranged between 60 and 100 μm, whereas the diameter of cell clumps with 2,4-D alone reached only 15–30 μm during the first 3 weeks of culture. Cell proliferation was also induced by treatment of root explants of RLD ecotype with NAA (3.0 mg l^{-1} α-naphthaleneacetic acid) and yielded higher regeneration rates in comparison with

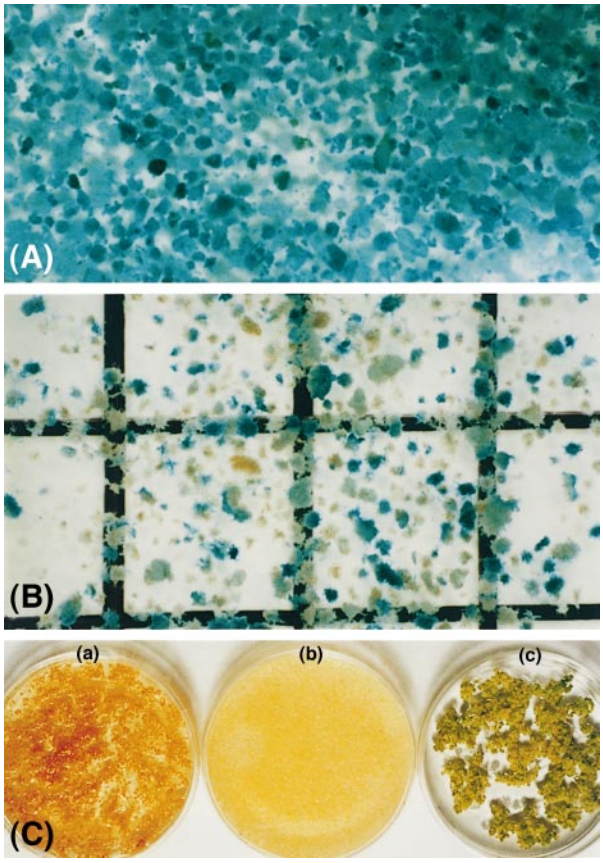

Figure 2. Transformation of *Arabidopsis* cell suspension with *Agrobacterium*.
(A) Histochemical GUS staining of *Arabidopsis* Col-1 cell suspension 3 days after co-cultivation with *Agrobacterium* carrying the vector pPCV6NFGUS.
(B) Counting GUS-stained cell clumps on a grid 21 days after co-cultivation with *Agrobacterium* carrying pPCV6NFGUS.
(C) Plate (a): *Arabidopsis* cell suspension co-cultivated with *Agrobacterium* GV3101 (pMP90RK) on CM containing hygromycin; plate (b): *Arabidopsis* cell suspension co-cultivated with *Agrobacterium* GV3101 (pMP90RK) carrying pPCV6NFGUS on CM with hygromycin; plate (c): induction of greening and shoot regeneration from transformed cell suspensions 18 days after plating on RM containing hygromycin.
The size of bar corresponds to 0.5 mm in (A) and (B), and to 2.0 cm in (C).

application of 2,4-D as sole auxin. Following an amplification phase of 5 weeks, all cell cultures initiated from *Arabidopsis* ecotypes Col-1, Col-5, RLD1 and WS2 were stabilized and showed a regular threefold increase in biomass during the subculture periods. Maintenance of a constant cell density (i.e. logarithmic growth) by regular sub-culturing was found to be essential, because high density (i.e. stationary) cultures formed larger cell clumps. The average size of microcalli was therefore controlled by sieving the cell cultures through a 250 μm mesh after every second subculture.

Agrobacterium *transformation of cell suspensions*

To transform the cells, *Agrobacterium* was added to the suspension cultures immediately after sub-culturing (see Experimental procedures). Bacterial attachment to plant cells (monitored by scanning electron microscopy, Figure 1d) occurred within 5–7 h and led to formation of long cellulose fibrils and pronounced clumping during co-culture for 48 h. Bacterial growth was arrested by claforan and tricarcillin-clavunic acid (both at 150 mg l^{-1}) to allow the plant cell cultures to proliferate without observable growth retardation until the end of the sub-culture period of 7 days.

The efficiency of *Agrobacterium* transformation was determined using the binary T-DNA vector pPCV6NFGUS, which was constructed by inserting a CaMV 35S promoter-driven *uidA* reporter gene, containing an intron (Vancanneyt *et al.*, 1990), as a *Hin*dIII fragment into the *Xba*I site of gene fusion vector pPCV6NFHyg (Figure 3, Koncz *et al.*, 1989). In addition to histological detection of the T-DNA-encoded β-glucuronidase (GUS) reporter exclusively in plant cells, T-DNA tags in plant genes were also identified by transformation with pPCV6NFGUS. This vector (as well as its precursor plasmid pPCV6NFHyg) carried a promoterless neomycin phosphotransferase (*aph(3′)II*) reporter gene linked to the right T-DNA border. Translational *aph(3′)II* gene fusions resulting from T-DNA integration into plant genes could thus be scored by selecting for kanamycin-resistant transformed calli. After *Agrobacterium* co-cultivation, aliquots of infected cell suspensions were plated on grids and stained with X-gluc (Figure 2A,B; Jefferson, 1987) to determine the frequency of cells displaying GUS activity. As summarized for the *Arabidopsis* ecotypes Col-1 and RLD1 in Table 1, the frequency of GUS-expressing cells decreased gradually from 93–95% to a stable value of about 50% between days 2 and 14 after co-cultivation, indicating that about half of the infected cells were stably transformed by T-DNA. Because not all cells showed GUS activity within the cell clumps, selection was applied against the untransformed cells using the T-DNA-encoded hygromycin resistance marker. As expected, about 50% of microcalli formed hygromycin-resistant colonies, whereas control cells co-cultivated with an *Agrobacterium* strain (GV3101 pMP90RK) lacking the vector pPCV6NFGUS turned brown and did not yield any viable colonies on hygromycin-containing CM (Figure 2C). Selection for kanamycin (50 or 100 mg l^{-1}) resistance after transformation with pPCV6NFGUS or pPCV6NFHyg yielded about 4–5 times fewer colonies in comparison with hygromycin selection (data not shown). This was in accordance with previous results showing that the average frequency of translational *aph(3′)II* gene fusions induced by T-DNA integration into plant genes ranges between 15 and 20% in *Arabidopsis* (Koncz *et al.*, 1989).

Plant regeneration

For high-frequency plant regeneration, only cultures that had undergone no more than 10–15 sub-culture periods

Table 1. *Agrobacterium* transformation and regeneration of *Arabidopsis* cell cultures

Days after co-cultivation	GUS-stained microcalli (%) *Arabidopsis* ecotypes	
	Col-1	RLD
2	98.38 ± 0.90	95.46 ± 0.41
3	84.14 ± 5.08	94.72 ± 1.96
4	69.56 ± 3.28	87.54 ± 9.53
5	66.81 ± 5.58	87.23 ± 5.19
6	68.56 ± 8.87	80.86 ± 13.08
7	68.76 ± 4.34	75.94 ± 12.57
14	52.10 ± 5.67	51.10 ± 1.74
21	52.98 ± 10.43	57.00 ± 9.94
Diameter of microcalli (μm)	Transformed colonies (%)	Regenerating colonies (%)
< 100	56.00 ± 0.70	11.19 ± 0.39
< 250	85.95 ± 1.76	15.92 ± 1.79
< 500	75.45 ± 0.91	51.15 ± 0.96
< 850	45.55 ± 0.49	81.30 ± 4.30
> 850	33.45 ± 3.18	93.17 ± 2.37

Upper section: Frequency of GUS-expressing cell clumps following co-cultivation with *Agrobacterium* GV3101 (pMP90RK) carrying pPCV6NFGUS.
Lower section: Correlation between the size of microcalli and frequencies of *Agrobacterium* transformation and plant regeneration. Frequency of GUS-stained cell clumps was scored 5 days after co-cultivation with *Agrobacterium*, whereas regeneration rate was observed 35 days after plating the microcalli on RM containing hygromycin.
Standard deviation of the data (±) resulted from three independent experiments.

proved to be suitable. The age of suspension cultures grown in the presence of 2,4-D dramatically affected both ploidy level and fertility of regenerated plants. From cell suspensions passing more than 40–50 sub-cultures, practically no fertile plants could be regenerated. Therefore, a careful standardization of the transformation protocol proved to be essential for obtaining continuously regenerating transformed callus and shoot cultures.

Plant regeneration was induced 7–14 days after co-cultivation with *Agrobacterium* by sub-culturing the cells in RM (regeneration medium, i.e. BM with 2.0 mg l^{-1} isopentenyl adenosine riboside (IPAR) and 0.05 mg l^{-1} NAA). Some ecotypes, such as RLD1 and Col-5, also showed regeneration in liquid RM. For clonal propagation, the cells were plated on solid RM containing hygromycin and antibiotics to control *Agrobacterium*. When kanamycin selection was applied, kanamycin-resistant calli were transferred after 3–4 sub-cultures to kanamycin-free medium to ensure high-frequency shoot regeneration. Transformed calli started to green (Figure 2C) within 15–20 days and yielded a mass of regenerating shoots after 30–40 days. Shoot explants were transferred to 0.5 BM with 0.5% sucrose in test tubes and produced a seed set, even without rooting, within 3–4 weeks.

To optimize the plant regeneration protocol, the transformed cell suspensions were either directly plated on RM, or embedded in gelrite layered on the surface of RM plates. Both plating techniques resulted in similar regeneration rates, but embedding in gelrite caused a delay of 7–14 days in shoot formation. Optimal (over 90%) regeneration frequencies were achieved when microcalli with an average size of 850 μm were plated on RM. In contrast, the highest transformation rates were obtained by *Agrobacterium* infection of microcalli with an average diameter of 250 μm. Therefore, to secure an optimal transformation efficiency, a size selection was carried out by sieving the suspensions through a mesh of 250 μm before co-cultivation with *Agrobacterium*, and the cultures were only plated on RM when the diameter of transformed microcalli reached about 850 μm (Table 1). This transformation protocol yielded a practically unlimited supply of transformed calli in each co-cultivation experiment. Because most transformed microcalli were capable of regenerating fertile plants, the yield of transgenic plants was only limited by the effort invested in plant propagation before the regeneration capability of cultures declined. A small proportion of plants obtained with pPCV6NFHyg transformation of *Arabidopsis* Col-1 was exploited to test the feasibility and efficiency of genome sequencing using random T-DNA tags.

Gene identification with sequenced T-DNA tags

To combine reverse genetics with genome sequencing, a protocol simplifying the rescue and sequencing of T-DNA-tagged chromosomal DNA segments was developed (see Experimental procedures). To isolate both left and right insert junctions of a T-DNA tag, plant DNAs are usually digested with an enzyme which does not cleave within the T-DNA and self-ligated as described by Koncz *et al.*, 1989, 1994). Plant DNA fragments entrapped by self-circularization between the T-DNA borders are isolated either by plasmid rescue (with the help of a plasmid replicon and a bacterial selectable marker carried by the T-DNA), or by PCR amplification and sequenced with primers facing the T-DNA ends. To double the output of genome sequencing, the plant DNA may also be digested with an enzyme which cleaves within the T-DNA tag. Following self-circularization, the plant DNA is thus ligated to known internal T-DNA sequences through an endonuclease cleavage site which may also serve as an RFLP marker. Following isolation by plasmid rescue or inverse PCR, the circularized plant DNA fragments can be bidirectionally sequenced using primers annealing to the left or right T-DNA arms. In addition to finding T-DNA-tagged genes, this sequencing method also allows the identification of plant genes that are located distantly from the ends of T-DNA tags. When two enzymes

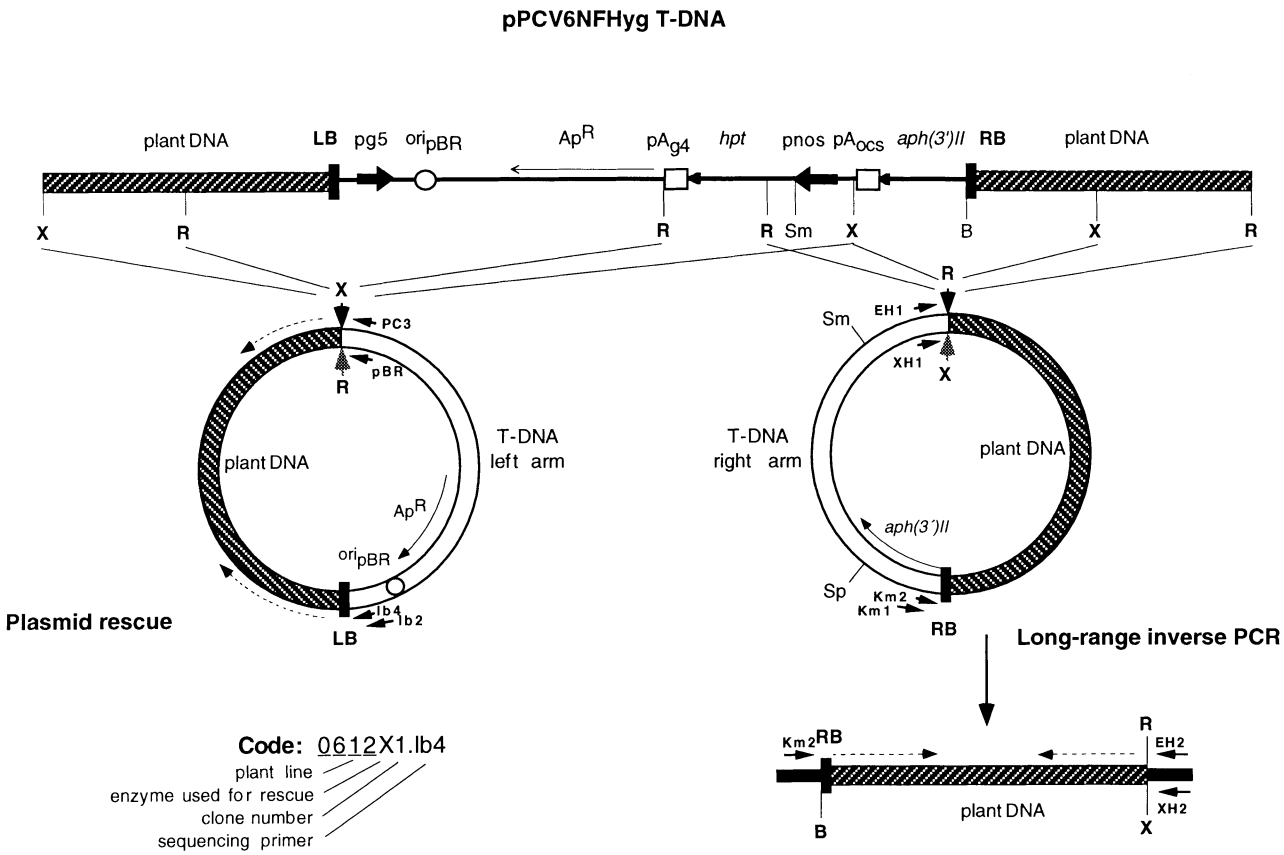

Figure 3. Genome sequencing with T-DNA tags.
A schematic map of pPCV6NFHyg T-DNA is shown at the top. To isolate the left T-DNA arm by plasmid rescue, the plant DNA is digested with *Eco*RI (R) or *Xba*I (X) and self-ligated before transformation into *E. coli* by electroporation. Plant DNA segments linked to the left T-DNA arm in the rescued plasmids are sequenced using primers lb2/lb4 at the left T-DNA 25 bp border (LB) and primers pBR and PC3 located upstream of the *Eco*RI and *Xba*I cleavage sites, respectively, within the T-DNA. The right T-DNA arm in junction with plant DNA sequences was similarly self-ligated after digestion with *Eco*RI or *Xba*I. The ligated plant DNA was linearized by *Sma*I (Sm) or *Sph*I (Sp) digestion, and used as template to amplify the plant DNA segment flanking the right T-DNA 25 bp border (RB) by long-range inverse PCR using primers at the right T-DNA border (Km1/Km2), as well as upstream of the *Eco*RI (EH1/EH2) and *Xba*I (XH1/XH2) sites within the T-DNA. The PCR products were gel-purified and sequenced with anchored primers Km1 and EH1 or XH1. Alternatively, the PCR fragments were digested with *Bam*HI (B) and *Eco*RI (R) or *Xba*I (X) and subcloned in pBluescript for sequencing. The code of sequenced T-DNA tags (left corner, Table 2) identifies the transformed plant line (e.g. 612), the enzyme used for isolation of a particular T-DNA tag (e.g. X1 = T-DNA insert no. 1 isolated with *Xba*I), and the primer used for sequencing (e.g. lb4, which also indicates whether the template was a plasmid or a PCR fragment).
Abbreviations: LB and RB, left and right T-DNA 25 bp border repeats; pg5, promoter of T-DNA gene 5; ori_{pBR}, pBR322 replication origin; Ap^R, ampicillin resistance gene; pA_{g4}, polyadenylation sequence of T-DNA gene 4; *hpt*, hygromycin phosphotransferase gene; pnos, promoter of the nopaline synthase gene; pA_{ocs}, polyadenylation sequence of the octopine synthase gene; *aph(3')II*, neomycin/kanamycin phosphotransferase gene.

are used for rescuing the T-DNA insert junctions, partial sequencing of up to six DNA segments (e.g. each with an average length of 450–650 bp, see below) may identify genes within a chromosomal locus of 4–10 kb. In addition, sequenced T-DNA tags provide RFLP/PCR and dominant selectable markers (e.g. hygromycin/kanamycin resistance) for genetic mapping as well as probes for physical mapping with yeast and bacterial artificial chromosome (YAC and BAC), or P1 phage clones.

To test the efficiency of this technique, DNA was isolated from F_2 progeny of 21 transgenic *Arabidopsis* (Col-1) plants carrying the T-DNA of pPCV6NFHyg gene fusion vector (Figure 3, the sequence of pPCV6NFHyg T-DNA is available from the Nottingham *Arabidopsis* Stock Center). Plant DNAs were digested independently with *Eco*RI or *Xba*I to cleave the T-DNA into two parts, and with *Hin*dIII which had no cleavage site within the T-DNA. Aliquots (1.0 μg) of digested DNAs were self-circularized by ligation. Half of the DNA samples (0.5 μg) were used for transformation of *E. coli* by electroporation to isolate plant DNA fragments linked to the left T-DNA arm carrying a pBR322 replicon and an ampicillin resistance gene. The rescued plasmids were sequenced with primers hybrizing to the ends of the left T-DNA arm (Figure 3, primers lb2/4, pBR or PC3). The second half of DNA samples was digested with *Sma*I or *Sph*I to cleave the right T-DNA arm and thereby linearize the self-circularized fragments. These DNAs served as templates for long-range PCR and sequencing with anchored T-DNA primers (Figure 3, primers Km1/2, EH1/2, Xh1/2) facing the termini of plant DNA fragments. Similarly,

Table 2. BLAST scores with sequenced T-DNA tags

T-DNA tagged sequence	Genbank accession number	Similarity to [accession number]
0612ea.lb	AF005781	None
0612xa.lb4	AF005782	Amine oxidase (copper-containing) [sp. P46881]
0612xa.pcr	AF005783	Copper amine oxidase [gb D38508]
0612xb.pcr	AF005784	None
0759xa.lb4	AF005785	None
0759xa.pcr	AF005786	None
0864e4.lb	AF005787	*Arabidopsis* ribosomal DNA spacer (#3) [emb X52636]
0864xc.lb2	AF005788	*Arabidopsis* repeat region (clone 164 A) [emb X92080]
0864xc.pcr	AF005789	*Arabidopsis CER3*-like gene [emb X95961]
0864xd.lb4	AF005790	Human Golgi protein LDLC [pir A53542]
0864xd.pcr	AF005791	None
084xdea.pbr	AF005792	Yeast MSP1 protein (TAT-binding homolog) [sp. P28737]
0941e1.lb4	AF005793	*Arabidopsis* cDNA [gb T45807], ASK1 kinase [sp. P43291]
0941e1.pbr	AF005794	*Arabidopsis* cDNA, Fe (II) transporter [gb N37871]
1562x2.lb4	AF005795	*Zea mays* pollen specific mRNA [gb I16762]
1772x3.lb4	AF005796	*Arabidopsis* farnesyl diphosphate synthase 1 [gb L46367]
1851xa.lb4	AF005797	None
1851xa.pcr	AF005798	None
1851xb.lb4	AF005799	None
1851xb.pcr	AF005800	None
1884xa.lb4	AF005801	None
1884xa.pcr	AF005802	None
2031e2.lb4	AF005803	None
2031e2.pbr	AF005804	*Arabidopsis* receptor-like protein kinase [gb M84660]
2046x4.lb4	AF005805	None
2051e3.lb4	AF005806	None
2051e3.pbr	AF005807	*Arabidopsis* cDNA clone 204G13T7 [gb H77156]
2276ea.lb	AF005808	None
2276xa.lb4	AF005809	None
2276xa.pcr	AF005810	None
2322e10.lb4	AF005811	None
2322e10.pbr	AF005812	*Arabidopsis* cDNA clone H6B5T7 [gb W43818]
2322e6.lb4	AF005813	None
2322e6.pbr	AF005814	*Brassica* mRNA for chitinase [emb X61488]
2454ec.lb4	AF005815	None
2454ec.pbr	AF005816	*Linum* mRNA for fis1 protein [emb X86733]
2454xa.lb4	AF005817	None
2454xa.pcr	AF005818	None
2454xaeb.pbr	AF005819	*Arabidopsis* 25S-18S ribosomal DNA spacer [emb X15550]
2462ea.lb4	AF005820	None
2462ea.pbr	AF005821	Tomato Pto kinase [gb U59316]
2462eb.lb4	AF005822	None
2462eb.pbr	AF005823	Yeast probable calcium-binding protein [sp. P36132]
2462xa.lb	AF005824	None
2734ea.lb4	AF005825	None
2734ea.pbr	AF005826	None
2761xa.lb4	AF005827	*Arabidopsis* mRNA for J-domain protein [emb Z49238]
2761xapcr.rev	AF005828	*Arabidopsis* mRNA for J-domain protein [emb Z49238]
2761xc.lb4	AF005829	*Arabidopsis* mRNA for J-domain protein [emb Z49238]
2761xc.pcr	AF005830	None
2981xae1.pbr	AF005839	*Arabidopsis* mRNA for RNA helicase [emb X97970]
2981xa.pcr	AF005840	*Arabidopsis* mRNA for RNA helicase [emb X97970]
2981xa.lb4	AF005841	*Arabidopsis* 81 kb genomic sequence [emb X98130]
2981pc.km	AF005831	*Arabidopsis* mRNA for unknown ORF02 [emb X97484]
3242e1.lb4	AF005832	*Arabidopsis* cDNA clone G5C10T7 [gb N96129]
3242e1.pbr	AF005833	None
3242e2.lb	AF005834	None
3322e1.lb4	AF005835	*Arabidopsis* cyc3b mRNA [emb Z31402]
3322e1.pbr	AF005836	None
Okae6.lb4	AF005837	None
Okae6.pbr	AF005838	*Zea mays* mRNA for porin [emb X73429]

Data libraries: gb, Genbank, sp., Swissprot, emb, EMBL.
Underlined: T-DNA tags in genes.

half of the *Hin*dIII-digested and self-circularized DNAs were transformed into *E. coli* to rescue plasmids, and the other half of the samples were linearized by *Sma*I and subjected to long-range PCR. The plasmid and PCR DNA templates were sequenced with primers facing the left and right T-DNA ends (Figure 3, Lb2/4 and Km1/2, respectively). Alternatively, the PCR fragments were subcloned in pBluescript before sequencing (see Experimental procedures).

The T-DNA rescue protocol was controlled by physical mapping of T-DNA insertions using Southern hybridization of plant DNAs (data not shown). Because about 40% of the T-DNA insertions were found in tandem repeats, the rescued plasmid clones and PCR fragments were always hybridized with T-DNA end-probes derived from the promoter region of gene 5 and the *aph(3′)II* gene of pPCV6NFHyg (Figure 3) to discard templates carrying self-joined T-DNA borders without plant DNA. Due to preferential use of the right 25 bp border sequence, about 10% of

the T-DNA inserts also contained an intact left border still linked to vector sequences located outside the T-DNA. These T-DNA inserts were also discarded after hybridization with a probe from the pPCV6NFHyg vector backbone (i.e. RK2 sequences; Koncz *et al.*, 1994) because they usually contained only short segments of plant DNA. In the remaining templates, the ends of plant DNA segments were automatically sequenced covering an average of 450–650 bp per reaction. The availability of overlapping *Eco*RI and *Xba*I clones facilitated the assembly of longer sequence segments, albeit at the cost of redundant sequencing. Thirty-four T-DNA-tagged chromosomal DNA fragments yielded total sequence information of 37 552 bp, which was subjected to BLASTX and BLASTN homology searches as described previously for the analysis of ESTs (Table 2; Höfte *et al.*, 1993; Newman *et al.*, 1994). Sequencing of 32 plant DNA templates with the T-DNA left border primers (Figure 3, lb2 and lb4) identified nine T-DNA-tagged genes. At a distance from the T-DNA tags, a significant sequence similarity to 16 Genbank accessions (genes, ESTs, proteins and BAC clones) was detected by sequencing of 29 templates with primers reading through endonuclease cleavage sites (Figure 3, primer pBR for *Eco*RI and primer PC3 for *Xba*I) which were used for self-circularization before plasmid rescue and LR-iPCR. Terminal sequences from 13 T-DNA-tagged plant DNA fragments shorter than 0.5 kb found no significant homology in the database.

Initial prospectives of genome sequencing by T-DNA tags

As expected from the comparison of estimated gene number and available unique ESTs in *Arabidopsis*, about one-third of sequenced T-DNA tags detected known *Arabidopsis* genes and ESTs (Table 2). Four of nine T-DNA inserts in genes were found in the promoter region of *Arabidopsis* genes encoding the ASK1 kinase (Swissprot P43291), cyclin 3b (EMBL accession number Z31402), ORF02 (EMBL accession number X97484) and EST mRNA (EMBL accession number N96129), respectively, whereas two tags were identified in the coding regions of a J-domain protein (EMBL accession number Z49238) and a farnesyl diphosphate synthase (FPS1, Genbank accession number L46367) gene. Furthermore, three T-DNA tags were localized in novel genes coding for orthologs of a copper amine oxidase, a peripheral Golgi-associated low density lipoprotein (LDLC, Genbank accession number U61947) and a pollen-specific maize mRNA (Genbank accession number I16762). From 16 sequences located at a distance from the T-DNA tags, only two showed nearly 100% identity with *Arabidopsis* ESTs, whereas a segment of identity with a common repeat element of 25S–18S ribosomal spacers was observed in another two sequences. Significant BLAST scores, but no identity, were detected between 12 sequenced *Arabidopsis* DNA segments and database accessions from plants and other organisms (Table 2). In summary, 62% of sequenced T-DNA tags scored significant similarity in the database, and half of these proved to be novel in *Arabidopsis*.

These data show that even a small set of sequenced T-DNA tags may yield a significant number of new genes which remained undetected by the EST sequencing project, as well as gene mutations that could not be identified so far using the available genetic and molecular methods. Further studies of some of these mutations may thus unravel the function of cyclin 3b, ASK1 kinase, LDLC or J-domain proteins in basic cellular processes. Although both plasmid rescue and LR-iPCR technologies proved to be equally efficient for the isolation of plant DNA junctions of any T-DNA tag (i.e. the frequency of obtaining no product was less than 1% using both methods), currently only the LR-iPCR method seems to be well adaptable to full automation. Our preliminary data indicate that, if appropriate controls were used to filter out PCR fragments resulting from amplification of tandem T-DNA insert junctions, up to 98% of PCR-amplified DNA fragments could yield suitable plant DNA sequence information, whereas the rest detect rare T-DNA insert rearrangements. Because technologies applied in current genome sequencing approaches are fully compatible with the PCR-based T-DNA end-sequencing approach (Huang and Miao, 1997; Miklos and Rubin, 1996), sequenced T-DNA tags may provide not only useful tools to extend the 'genes galore' accumulated by ESTs (Cooke *et al.*, 1996; Höfte *et al.*, 1993; Newman *et al.*, 1994; Sasaki *et al.*, 1994), but also to elucidate the function of plant genes by reverse genetics.

Experimental procedures

Culture media

One litre of basal medium (BM, pH 5.8) contained 4.33 g of MS inorganic salts (Murashige and Skoog, 1962; Sigma M-5524), 3.0% sucrose and 2× B5 vitamins (Gamborg *et al.*, 1968), including 2 mg nicotinic acid, 2 mg pyridoxin HCl, 20 mg thiamine–HCl, and 200 mg *myo*-inositol. Gelrite 0.2% (Phytagel, Sigma) was used as gelling agent in BM, CM and RM. Seeds were germinated in 0.5 × MS containing half the concentration of MS inorganic salts, 0.5% sucrose and 0.8% agar (Koncz *et al.*, 1994). Root cultures were maintained in callus medium (CM) containing 0.5 mg 2,4-dichlorophenoxyacetic acid (2,4-D), 2.0 mg indole-3-acetic acid (IAA), and 0.5 mg 6-(γ,γ- dimethylallylamino)-purine riboside (IPAR) in 1 litre of BM. Shoots were induced in regeneration medium (RM), consisting of BM with 2.0 mg l^{-1} IPAR and 0.05 mg l^{-1} α-naphtaleneacetic acid (NAA), and transferred to test tubes containing 0.5× MS solidified with gelrite to obtain seeds. *Agrobacterium* media were as described by Koncz *et al.*, 1994.

Establishment and transformation of Arabidopsis *cell suspensions*

About 5000 seeds (0.1 g) of each of *Arabidopsis* ecotypes Col-1, Col-5 (also known as C24), WS2 and RLD1 (Ohio *Arabidopsis*

Biological Resource Center, seeds@genes ys.cps. msu.edu) were surface sterilized in Eppendorf tubes with 1 ml of 5% calcium hypochlorite containing 0.02% Triton X-100 (or 10% (v/v) sodium hypochlorite, 0.1% Triton X-100) for 15 min, pelleted by centrifugation, washed five times with 1 ml of sterile water, dried, and germinated on 0.5× MS agar plates exposed to a 16 h light/8 h dark period at 25°C. Fifteen to twenty one-week-old seedlings were placed in an Erlenmeyer flask containing 35 ml of liquid BM and cultured on a rotary shaker (120 rev min^{-1}) for 15–20 days under similar conditions. Roots of seedlings (⊤ 3 g) harvested from a flask were dissected, cut in pieces (⊤ 2 mm), transferred into 50 ml of liquid CM, and cultured in the dark for 15–21 days. Alternatively, approximately 1000 seeds were directly germinated in the dark in 50 ml of CM to initiate cell suspensions from seedling-derived calli in a similar fashion. Cell clumps released from root or seedling explants were separated by filtration through a sieve of 850 μm (or 500 μm). Root explants were returned to CM to generate more cells in suspension, and the filtered cells were pelleted, suspended in 30–50 ml of fresh CM, and allowed to proliferate for 7 days of regular sub-culture. During the first two sub-cultures, cells from a single flask were dispensed into two flasks of 250 ml, pelleted, and after removing the excess of medium were resuspended in 50 ml of fresh CM. During subsequent sub-cultures, 50 ml CM was added to each flask, and the resulting 100 ml of culture was equally dispensed into three flasks. Using this strategy, a fine cell suspension was obtained from each *Arabidopsis* ecotype after 5 weeks. The size of cell clumps was controlled by filtering the cultures through a sieve of 250 μm at every second sub-culture. Logarithmic growth of cultures was controlled by monitoring the increase of fresh weight as described (Street, 1973).

An aliquot of 1.5 ml of *Agrobacterium* culture (OD_{550}: 1.0) grown as described (Koncz *et al.*, 1994) was pelleted in an Eppendorf tube by centrifugation, resuspended in 1.5 ml CM, and added to 35 ml of cell suspension at the time of sub-culturing. The co-cultivation was carried out for 48 h, then claforan and tricarcillin-clavulanic acid (150 mg l^{-1} of each) were added to kill *Agrobacterium*. The cells were harvested 5 days later, and either sub-cultured for another 7 days, or directly plated on RM. Cells from *Agrobacterium* infected cultures were collected in a 50 ml Flacon tube by centrifugation at 1000 ***g*** for 3 min, washed twice with 3.0% sucrose, and counted on a 2 cm^2 grid. Cells suspended in liquid RM at a density of 10^3 cells ml^{-1} were plated on RM with gelrite containing antibiotics for selection of transformed cells (hygromycin 15 mg l^{-1} or kanamycin 100 mg ml^{-1}) and to control *Agrobacterium*. Alternatively, cells were suspended in 0.2% gelrite dissolved in 3.0% sucrose solution (pH 5.8), and layered on the top of RM plates. The regenerating calli were cultured under continuous light at 22°C until green colonies appeared, which were counted after 35 days by placing a 2 cm^2 grid on the Petri dishes. Colonies in at least three randomly chosen squares from ten plates per experiment were counted. Regenerated shoots were individually transferred to test tubes (length: 20 cm; diameter: 3 cm) containing 10–15 ml of 0.5× MS with gelrite, closed with loose cotton plugs, and cultured for 6–7 weeks to set seeds at 25°C using a 16 h light/8 h dark period.

Agrobacterium *vectors and histochemical GUS-staining*

Plasmid pPCV6NFGUS was constructed by cloning of a *Hin*dIII fragment of p35SGUSINT from pBIN19 (Vancanneyt *et al.*, 1990) into the *Xba*I site of pPCV6NFHyg (Koncz *et al.*, 1989) after filling-in the ends of DNA fragments with T4 DNA polymerase to regenerate the *Xba*I site (Sambrook *et al.*, 1989). In addition to known features of pPCV6NFHyg (Figure 3; Koncz *et al.*, 1989), pPCV6NFGUS carried a *uidA* (β-glucuronidase, GUS) reporter gene with a portable intron from the potato *ST-LS1* gene under the control of the 35S RNA promoter of cauliflower mosaic virus (CaMV 35S). This *uidA* reporter gene therefore yielded no detectable GUS activity in *Agrobacterium* (Vancanneyt *et al.*, 1990). Transfer of pPCV6NFGUS and pPCV6NFHyg from *E. coli* to *Agrobacterium* GV3101 (pMP90RK), and the use of *Agrobacterium* strains for plant transformation were as described by Koncz *et al.*, 1994. To monitor the transformation rates with pPCV6NFGUS, aliquots of cell cultures (0.5 ml) sampled at various times after *Agrobacterium* co-cultivation were stained with X-gluc (1 mg ml^{-1} X-gluc in 50 mM sodium phosphate buffer (pH 7.0) containing 0.5 mM potassium ferricyanide and 0.5 mM potassium ferrocyanide) for 6 h, then plated on grids to count the number of GUS-expressing microcalli in at least three squares of 1 cm^2 using ten samples from each experiment. As control, changes in GUS activity were also assayed after co-cultivation by fluorimetric enzyme assays (Jefferson, 1987).

Isolation of T-DNA tagged chromosomal DNA fragments by plasmid rescue

Plants grown aseptically in 0.5× MS containing hygromycin (15 mg ml^{-1}) were harvested prior to flowering to purify DNA on a large scale using CsCl-banding or on a mini-scale with or without CTAB precipitation (Dellaporta *et al.*, 1983; Taylor and Powell, 1983). Other methods yielding high molecular weight genomic DNA free of protein and RNA (e.g. Souer *et al.*, 1995; Krysan *et al.*, 1996) were also found to be applicable. An aliquot of 5 μg of total plant DNA was digested with 100 U of restriction endonucleases *Eco*RI, *Xba*I or *Hin*dIII for at least 6 h at 37°C in a volume of 200 μl of enzyme buffer (Sambrook *et al.*, 1989). After testing 10 μl aliquots by agarose gel electrophoresis, the samples were phenol/chloroform extracted and precipitated with i-propanol (Sambrook *et al.*, 1989).

To isolate self-circularized plant fragments by T-DNA-mediated plasmid rescue, 1 μg of digested DNA samples were ligated in a volume of 200 μl ligase buffer (Sambrook *et al.*, 1989) containing 10% polyethylene glycol (PEG 4000) at 15°C for 8 h. The samples were phenol/chloroform extracted, precipitated by i-propanol, washed twice by 70% ethanol, dried, and dissolved in water before electroporation into *E. coli* DH1 cells (Zabarovsky and Weinberg, 1990) using a BioRad Gene Pulser (settings 2.5 kV, pulse controller 200 Ω, (capacitance extender 250 μFD for a 0.2 ml BioRad cuvette). Plasmids carrying a segment of a tandem T-DNA repeat or intact left T-DNA border were identified by colony hybridization (Sambrook *et al.*, 1989) using ^{32}P-labelled probes of *Xba*I–*Bam*HI and *Xho*I–*Kpn*I fragments from pPCV6NFHyg which contained sequences from the right border-linked *aph(3′)II* gene and vector backbone outside of the T-DNA, respectively (Koncz *et al.*, 1994). All rescued plasmids were subjected to physical mapping, and sequenced using a PCR kit (ABI Prism^{-} Dye Terminator Cycle Sequencing) and an automatic sequencer (ABI 377). The positions of sequencing primers lb2 (5′-GACCCTTACCGCTTTAGTT-CCGTAGCTAGCACTTC-3′) and lb4 (5′-AGAGGTATAACTGGTAG-TATGAG-3′) facing the left T-DNA border, as well as that of primer pBR (5′-CCTATAAAAATAGGCGTATCACGAGGCCC-3′) at the *Eco*RI site and primer PC3 (5′-CCTTGCGCCCTGAGTGCTTGCGGCAGC-3′) at the *Xba*I site are shown in Figure 3. *Hin*dIII fragments of plant DNA rescued as plasmids by self-circularization were sequenced with the lb2/lb4 and Km1/Km2 primers (see below).

Amplification of T-DNA tagged plant DNA fragments by LR-iPCR

Plant DNA sequences flanking the right arm of pPCV6NFHyg T-DNA were isolated by long-range inverse PCR (LR-iPCR). After digestion with *Eco*RI or *Xba*I and self-circularization by ligation, 0.5 μg of plant DNA was digested with either *Sma*I or *Sph*I to linearize the DNA templates by cleaving the right T-DNA arm (see Figure 3). These DNA samples were subjected to PCR amplification using primer Km1 (5′-CAAAGCGAACCACCAGCTTACCCGTCCAT-CGGC-3′) facing the T-DNA right border, and either primer EH1 (5′-TCCTGCGGGTAAATAGCTGCGCCGATGG-3′) at the *Eco*RI site, or primer XH1 (5′-TAGCTCAGATCCTTACCGCCGGTTTCGG-3′) at the *Xba*I site. PCR reactions were performed in 50 μl using either Elongase PCR (BRL) or LA-PCR (Takara Shuzo Co.) kits as recommended by the suppliers. DNA samples were denatured at 95°C for 2 min, and amplified using 35 cycles (94°C for 30 sec, 65°C for 30 sec, 68°C for 8 min) followed by elongation at 68°C for 10 min. The PCR products were resolved on agarose gels and isolated (Sambrook *et al.*, 1989). When no amplified DNA fragment was detected, a second PCR amplification was performed using 1 μl from a 500-fold diluted first PCR reaction mixture in combination with the nested primers Km2 (5′-CAGTCATAGCCGAATAGCCTCT-CCACCC-3′), and EH2 (5′-CGTTATGTTTATCGGCACTTTGCATCGG-3′), or XH2 (5′-CCGTTCAATTTACTGATTGTCCAAGCTC-3′). The isolated PCR fragments were either used directly as templates for sequencing, or digested with *Bam*HI and *Eco*RI or *Xba*I (see Figure 3) for subcloning in pBluescript (Stratagene) before sequencing. Sequence analyses were performed using the GCG and BLAST computer program packages as described for Genbank database searches with ESTs (Newman *et al.*, 1994).

Acknowledgements

This work was supported as part of a joint project between the Max-Planck Institut (Köln) and the Biological Research Center of the Hungarian Academy of Sciences (Szeged) by the Deutsche Forschungsgemeinschaft (DFG), as well as by the European Communities Biotech Programme BIO4-CT95–0183 '*Arabidopsis* insertional mutagenesis for the functional analysis of sequences identified by ESSA (AIM)', and OTKA grant T13182.

References

An, G. (1985) High efficiency transformation of cultured tobacco cells. *Plant Physiol.* **79**, 568–570.

Ballinger, D.G. and Benzer, S. (1989) Targeted gene mutations in *Drosophila. Proc. Natl Acad. Sci. USA*, **86**, 9402–9406.

Bancroft, I., Jones, J.D. and Dean, C. (1993) Heterologous transposon tagging of the *DRL1* locus in *Arabidopsis. Plant Cell*, **5**, 631–638.

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

Cardon, G., Frey, M., Saedler, H. and Gierl, A. (1993) Mobility of the maize transposable element *En/Spm* in *Arabidopsis thaliana. Plant J.* **3**, 773–784.

Cooke, R., Raynal, M., Laudié, M. *et al.* (1996) Further progress towards a catalogue of all *Arabidopsis* genes: analysis of a set of 5000 non-redundant ESTs. *Plant J.* **9**, 101–124.

Damm, B.R., Schmidt, R. and Willmitzer, L. (1989) Efficient transformation of *Arabidopsis thaliana* using direct gene transfer to protoplasts. *Mol. Gen. Genet.* **217**, 6–12.

Dellaporta, S.L., Wood, J. and Hicks, J.B. (1983) A plant DNA minipreparation: version II. *Plant Mol. Biol. Rep.* **1**, 19–21.

Feldmann, K.A., Malmberg, R. and Dean, C. (1994) Mutagenesis in *Arabidopsis*. In *Arabidopsis* (Meyerowitz, E.M. and Somerville, C.R., eds). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 137–172.

Forsthoefel, N.R., Wu, Y., Schulz, B., Bennett, M.J. and Feldmann, K.A. (1992) T-DNA insertion mutagenesis in *Arabidopsis*: prospects and perspectives. *Aust. J. Plant Physiol.* **19**, 353–366.

Franzmann, L.H., Yoon, E.S. and Meinke, D.W. (1995) Saturating the genetic map of *Arabidopsis thaliana* with embryonic mutations. *Plant J.* **7**, 341–350.

Gamborg, O.L., Miller, R.A. and Ojima, K. (1968) Nutrient requirements of suspension cultures of soybean root cells. *Exp. Cell Res.* **50**, 151–158.

Hayashi, H., Czaja, I., Schell, J. and Walden, R. (1992) Activation of plant gene by T-DNA tagging: auxin independent growth *in vitro. Science*, **258**, 1350–1353.

Hiei, Y., Ohta, S., Komari, T. and Kumashiro, T. (1994) Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. *Plant J.* **6**, 271–282.

Höfte, H., Desprez, T., Amselem, J. *et al.* (1993) An inventory of 1152 expressed sequence tags obtained by partial sequencing of cDNAs from *Arabidopsis thaliana. Plant J.* **4**, 1051–1061.

Horsch, R.B., Fraley, R., Rogers, S., Fry, J., Klee, H., Shah, D., McCormick, S., Niedermeyer, J. and Hoffmann, N. (1987) *Agrobacterium*-mediated transformation of plants. In *Plant Tissue and Cell Culture* (Green, C.E., Sommers, D.A., Hackett, W.P. and Biersboer, D.D., eds). New York: Alan Liss Inc., pp. 317–329.

Huang, G.M. and Miao, G.-H. (1997) The dawn of the post-genome era, seen from the ocean front. *Trends Biotechnol.* **15**, 200–202.

Jefferson, R.A. (1987) Assaying chimeric genes in plants: the GUS gene fusion system. *Plant Mol. Biol. Rep.* **5**, 387–405.

Kaiser, K. and Goodwin, S.F. (1990) 'Site-selected' transposon mutagenesis of *Drosophila. Proc. Natl Acad. Sci. USA*, **87**, 1686–1690.

Kakimoto, T. (1996) *CKI1*, a histidine kinase homolog implicated in cytokinin signal transduction. *Science*, **274**, 982–985.

Kertbundit, S., De Greve, H., Deboeck, F., Van Montagu, M. and Hernalsteens, J. (1991) *In vivo* random β-glucoronidase gene fusions in *Arabidopsis thaliana. Proc. Natl Acad. Sci. USA*, **88**, 5212–5216.

Knapp, S., Larondelle, Y., Rossberg, M., Furtek, D. and Theres, K. (1994) Transgenic tomato lines containing *Ds* elements at defined genomic positions as tools for targeted transposon tagging. *Mol. Gen. Genet.* **243**, 666–673.

Koes, R., Souer, E., van Houwelingen, A. *et al.* (1995) Targeted gene inactivation in petunia by PCR-based selection of transposon insertion mutants. *Proc. Natl Acad. Sci. USA*, **92**, 8149–8153.

Kohchi, K., Fujishige, K. and Ohyama, K. (1995) Construction of equalized cDNA library from *Arabidopsis thaliana. Plant J.* **8**, 771–776.

Koncz, C., Martini, N., Mayerhofer, R., Koncz-Kálmán, Z., Körber, H., Rédei, G.P. and Schell, J. (1989) High-frequency T-DNA-mediated gene tagging in plants. *Proc. Natl Acad. Sci. USA*, **86**, 8467–8471.

Koncz, C., Martini, N., Szabados, L., Hrouda, M., Bachmair, A. and Schell, J. (1994) Specialized vectors for gene tagging and expression studies. In *Plant Molecular Biology Manual*, Volume B2 (Gelvin, S.B. and Schilperoort, R.A., eds). Dordrecht, The Netherlands: Kluwer Academic Press, pp. 1–22.

Koncz, C. and Rédei, G.P. (1994) Genetic studies with *Arabidopsis*: a historical view. In *Arabidopsis* (Meyerowitz, E.M. and

Somerville, C.R., eds). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 223–252.

Krysan, P.J., Young, J.C., Tax, F. and Sussman, M.R. (1996) Identification of transferred DNA insertions within *Arabidopsis* genes involved in signal transduction and ion transport. *Proc. Natl Acad. Sci. USA*, **93**, 8145–8150.

Long, D., Martin, M., Sundberg, E., Swimburne, J., Puangsomlee, P. and Coupland, G. (1993) The maize transposable element system *Ac/Ds* as a mutagen in *Arabidopsis*: identification of an albino mutation induced by *Ds* insertion. *Proc. Natl Acad. Sci. USA*, **90**, 10370–10374.

Mathur, J., Szabados, L. and Koncz, C. (1995) A simple method for isolation, liquid culture, transformation and regeneration of *Arabidopsis thaliana* protoplasts. *Plant Cell Rep.* **14**, 221–226.

McKinney, E.C., Ali, N., Traut, A., Feldmann, K.A., Belostotsky, D.A., McDowell, J.M. and Meagher, R.B. (1995) Sequence-based identification of T-DNA insertion mutations in *Arabidopsis*: actin mutants *act2–1* and *act4–1*. *Plant J.* **8**, 613–622.

Meyerowitz, E.M. (1994) Structure and organization of the *Arabidopsis thaliana* nuclear genome. In *Arabidopsis* (Meyerowitz, E.M. and Somerville, C.R., eds). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 21–36.

Miao, Z.-H. and Lam, E. (1995) Targeted disruption of the *TGA3* locus in *Arabidopsis thaliana*. *Plant J.* **7**, 359–365.

Miklos, G.L.G. and Rubin, G.M. (1996) The role of the genome project in determining gene function: insights from model organisms. *Cell*, **86**, 521–529.

Morris, P.C. and Altmann, T. (1994) Tissue culture and transformation, In *Arabidopsis* (Meyerowitz, E.M. and Somerville, C.R., eds). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 173–222.

Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant.* **15**, 473–497.

Newman, T., de Bruijn, F.J., Green, P. *et al.* (1994) Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones. *Plant Physiol.* **106**, 1241–1255.

Osborn, B.I., Corr, C.A., Prince, J.P., Hehl, R., Tanksley, S.D., McCormick, S. and Baker, B. (1991) *Ac* transposition from a T-DNA can generate linked and unlinked clusters of insertions in the tomato genome. *Genetics*, **129**, 833–844.

Rushforth, A.M., Saari, B. and Anderson, P. (1993) Site-selected insertion of the transposon Tc1 into a *Caenorhabditis elegans* myosin light chain gene. *Mol. Cell Biol.* **13**, 902–910.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Sasaki, T., Song, J., Koga-Ban, Y. *et al.* (1994) Toward cataloguing all rice genes: large-scale sequencing of randomly chosen rice cDNAs from a callus cDNA library. *Plant J.* **6**, 615–624.

Schmidt, R., West, J., Love, K., Lenehan, Z., Lister, C., Thompson, H., Bouchez, D. and Dean, C. (1995) Physical map and organization of *Arabidopsis thaliana* chromosome 4. *Science*, **270**, 480–483.

Souer, E., Quattrocchio, F., de Vetten, N., Mol, J. and Koes, R. (1995) A general method to isolate genes tagged by a high copy number transposable element. *Plant J.* **7**, 677–685.

Street, H.E. (1973) Cell (suspension) cultures – techniques. In *Plant Tissue and Cell Culture* (Street, H.., ed.). Oxford: Blackwell, pp. 59–99.

Sunderesan, V., Springer, P., Volpe, T., Hawrd, S., Jones, J.D.G., Dean, C., Ma, H. and Martienssen, R. (1995) Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable element. *Genes Devel.* **9**, 1797–1810.

Taylor, B. and Powell, A. (1983) Isolation of plant DNA and RNA. *BRL Focus*, **4**, 4–6.

Topping, J.F., Wei, W. and Lindsey, K. (1991) Functional tagging of regulatory elements in the plant genome. *Development*, **112**, 1009–1019.

Vancanneyt, G., Schmidt, R., O'Connor-Sanchez, A., Willmitzer, L. and Rocha-Sosa, M. (1990) Construction of an intron-containing marker gene: splicing of the intron in transgenic plants and its use in monitoring early events of *Agrobacterium*-mediated plant transformation. *Mol. Gen. Genet.* **220**, 245–250.

Zabarovsky, E.R. and Weinberg, G. (1990) High efficiency electroporation of ligated DNA into bacteria. *Nucl. Acids Res.*, **18**, 5912.

Zachgo, E.A., Wang, M.L., Dewdney, J., Bouchez, D., Camilleri, C., Belmonte, S., Huang, L., Dolan, M. and Goodman, H.M. (1996) A physical map of chromosome 2 of *Arabidopsis thaliana*. *Genome Res.* **6**, 19–25.

Zwaal, R.R., Broeks, A., van Meurs, J., Groenen, J.T.M. and Plasterk, R.H.A. (1993) Target-selected gene inactivation in *Caenorhabditis elegans* using a frozen transposon insertion mutant bank. *Proc. Natl Acad. Sci. USA*, **90**, 7431–7435.