7

Transposon and T-DNA mutagenesis

MARK G. M. AARTS, CSABA KONCZ and ANDY PEREIRA

1. Introduction

Insertional mutagenesis with a known mobile DNA insert can generate mutations that are marked by a molecular tag. When the insertion disrupts a gene, causing a mutant phenotype, the tagged mutant gene can easily be cloned using the DNA insert as a molecular probe. DNA sequences of the cloned mutant gene can then be used to isolate the corresponding wild-type allele. This procedure, known as gene tagging, facilitates the analysis of both mutant and gene.

In higher organisms the gene tagging technique was first used in *Drosophila*, with well studied transposable elements as tags (1). Although transposable elements were long since known to be present in plants, the active use of plant transposons, for gene tagging purposes, first required their molecular characterization. Both maize and snapdragon transposons subsequently became standard tools for gene isolation (2).

The success of transposon tagging encouraged the development of heterologous tagging systems, using maize transposons in species lacking well characterized endogenous transposons, such as tobacco and *Arabidopsis thaliana* (3). At the same time the T-DNA, a unique insertion element transferred by *Agrobacterium tumefaciens* into plants, appeared to be equally suited for use as a molecular tag. Mutants caused by T-DNA insertion were found after generating sufficient numbers of T-DNA transformants and they were used for the isolation of genes in *Arabidopsis* (4–6; see also Chapter 6).

In this chapter various transposon and T-DNA tagging systems for *Arabidopsis* are described including strategies for their use in mutant analysis and gene isolation. *Arabidopsis* has a special advantage over other species for gene tagging, apart from being a model organism, as it is relatively easy to grow large populations in a very small area.

2. Transposon tagging

2.1 Endogenous transposable elements

Transposable elements (7) have been discovered and studied in detail by Barbara McClintock, who attributed the genetic instability of certain maize traits to genetic elements which are able to change their location (i.e. transpose) within the genome. Elements which transpose on their own are called autonomous elements, in contrast to non-autonomous elements (often deletion derivatives of autonomous elements), which require the presence of the autonomous elements for mobility. *Activator* (*Ac*), an autonomous element discovered by McClintock, is capable of activating a family of non-autonomous *Dissociation* (*Ds*) elements. *Suppressor-mutator* (*Spm*), another autonomous element, can activate non-autonomous *defective Spm* (*dSpm*) elements. Parallel to the discovery of *Spm* an autonomous element called *Enhancer* (*En*) was described, and found to activate non-autonomous *Inhibitor* (*I*) elements (8). After isolation and DNA sequencing of both the *En* and *Spm* elements from maize (9, 10) they turned out to be virtually identical, none the less both names are still used.

Ac and *En/Spm* both encode transposase proteins and contain short terminal inverted repeats. Additional subterminal regions with short inverted and direct repeats are required for transposition. *Ac* and *En/Spm* belong to different transposon families, which cannot activate each other. Soon after molecular characterization of *Ac* and *En/Spm* their extraordinary use in gene isolation was demonstrated by the cloning of transposon tagged genes encoding various steps of the anthocyanin biosynthesis pathway in maize. The mutation frequencies at target loci observed with *Ac* and *En/Spm* in maize are about 1–10 × 10^{-6} (2).

Although ubiquitous in plants, transposable elements from only a few species other than maize (*Antirrhinum majus*, *Petunia hybrida*) are characterized in sufficient detail to allow their use in gene tagging (11, 12). A poor knowledge of endogenous transposons is mainly the reason why maize transposons were exploited to design heterologous transposon tagging systems. Experiments with heterologous systems were initially carried out with *Ac* and *En/Spm* in tobacco (reviewed in ref. 3), followed a little later in *Arabidopsis*. Whether the recently characterized active endogenous transposon of *Arabidopsis* (13) will replace the use of available heterologous transposon tagging systems remains to be seen.

2.2 Transposon tagging systems in *Arabidopsis*

Transposon tagging in heterologous species can be employed using either one or two element systems. In a one element system, an autonomous transposable element is used as a mutagen. In a two element system, a non-autonomous transposable element is used, activated *in cis* or *trans* by the expression of a transposase, e.g. from a stable derivative of an autonomous element (14). In either case, the transposon is cloned into a selectable marker gene (e.g. for antibiotic resistance), blocking its expression (*Figure 1*). Excision of the element restores the activity of the excision marker gene, which can be followed with a selective agent (e.g. an antibiotic) in a phenotypic excision assay (15). In *Arabidopsis* both one and two element systems have been developed from the maize *Ac-Ds* and *En/Spm-I/dSpm* elements.

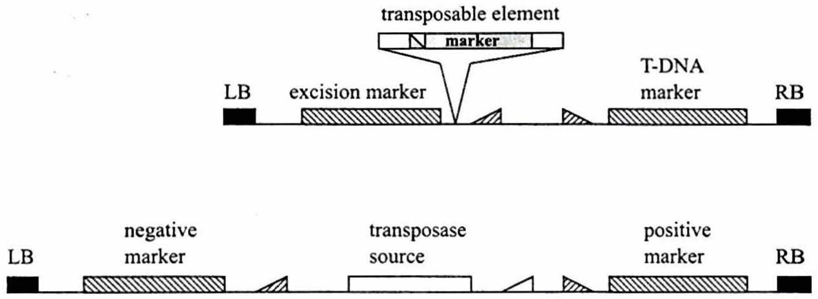

Figure 1. Basic design of a heterologous two element transposon tagging system. A marked T-DNA (*top*), contains a marked non-autonomous transposable element, inserted in an excision marker gene. Transposition of the transposable element occurs in the presence (*in cis* or *in trans*) of a T-DNA containing a transposase source with accompanying positive and negative selection markers. Triangles represent promoter sequences, LB and RB are left and right T-DNA borders respectively.

2.2.1 *Ac-Ds* systems

The unaltered *Ac* element from maize is not very active in *Arabidopsis*. The germinal excision frequency (the fraction of seedlings in which excision has occurred among the total number of seedlings in the progeny of a plant with *Ac*) of 0.2–0.5% is insufficient for efficient gene tagging (16, 17). Deletion of a methylation-sensitive CpG-rich *Nae*I fragment from the 5′ end of *Ac* increased the germinal excision frequency to a level suitable for gene tagging (18).

More efficient transposon tagging systems were designed based on two element systems. *Ds* elements, carrying a selectable marker, were mobilized by stable transposase sources yielding different germinal excision frequencies depending on the strength and timing of the promoter fused to the transposase gene (18–22). Frequencies of over 30% were achieved using an *Ac* transposase construct driven by the CaMV 35S promoter (20–22). This high frequency of excision was not as advantageous as expected. Over-expression of *Ac* transposase in tobacco inhibited late transposition of *Ds* elements (23), thus a predominantly early transposition yielded only a few different *Ds* inserts in the progeny (24).

2.2.2 Tagging with *Ac-Ds*

Many *Ac-Ds* two element systems in *Arabidopsis* have been published as summarized in *Table 1*. Although only tested in a few cases, the *Ac* transposase producing lines and the *Ds* lines can easily be combined from different systems, provided the selectable markers are compatible. A tagging strategy

Table 1. *Ac-Ds* two element systems

	Stable *Ac* transposase sources		*Ds* elements			
Transposase construct[a]	**Positive selection**	**Negative selection**[b]	**T-DNA selection**	***Ds* excision selection**	***Ds* element selection**	**Ref.**
wt*Ac*3′Δ	Kanamycin	X-Gluc	Kanamycin	Streptomycin	None	18
Δ*NaeIsAc*	Kanamycin	X-Gluc				18
Δ*NaeIsAc*	Kanamycin	NAM				18
35S-*Ac*	Hygromycin	None	Hygromycin	Kanamycin	35S readout	19
			None	Kanamycin	Methotrexate	19
p*Ac*-*Ac*	Kanamycin	None	Kanamycin	Streptomycin	None	20, 24
ocs-*Ac*	Kanamycin	X-Gluc	Kanamycin	Streptomycin	Hygromycin,	20, 24
35S-*Ac*	Kanamycin	None			35S readout	20, 24
35S-*Ac*	Methotrexate	X-Gluc	Hygromycin	Kanamycin	Chlorsulfuron	22
rbcS-*Ac*	Methotrexate	X-Gluc				22
chs-*Ac*	Methotrexate	X-Gluc				22
35S-*Ac*	Kanamycin	X-Gluc	Kanamycin	Chlorsulfuron	hygromycin,	21
35S-Δ*NaeIAc*	Kanamycin	X-Gluc			promoter or enhancer trap	21
wt*Ac*3′Δ	Hygromycin	None	Hygromycin	Kanamycin	BASTA®	54
35S-*Ac*	Kanamycin	NAM	Kanamycin, NAM	Selection for transposition	Kanamycin, enhancer/gene trap	45

[a] Key to the constructs:
wt*Ac*3′Δ: wild-type *Ac* element with 3′ terminal deletion.
Δ*NaeIsAc*: wild-type *Ac* element with 5′ *NaeI* fragment deletion.
35S-*Ac*: CaMV 35S promoter fused to *Ac* transposase gene.
p*Ac*-*Ac*: *Ac* promoter fused to *Ac* transposase gene.
ocs-*Ac*: octopine synthase promoter fused to *Ac* transposase gene.
rbcS-*Ac*: ribulose 1,5-biphosphate carboxylase small subunit promoter fused to *Ac* transposase gene.
chs-*Ac*: chalcone synthase promoter fused to *Ac* transposase gene.
35S-Δ*NaeIAc*: CaMV 35S promoter fused to *Ac* transposase gene with *NaeI* fragment deletion.
[b] X-Gluc: *p*-nitrophenyl β-D-glucuronide. NAM: α-naphthalene-acetamide.

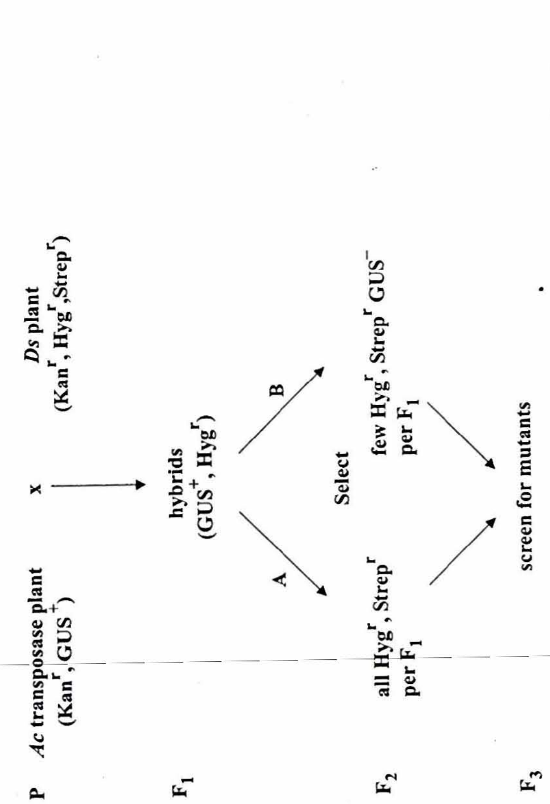

Figure 2. Strategy for tagging with two element *Ac-Ds* transposon systems in *Arabidopsis thaliana*. Strategy A is used with the Δ*NaeIsAc* transposase source. Strategy B is used with the CaMV 35S-*Ac* transposase source. The parental plants are both homozygous for the T-DNAs containing the *Ac* transposase source marked with β-glucuronidase (GUS^+), or the *Ds* marked with hygromycin resistance (hyg^r), inserted in the streptomycin resistance ($strep^r$) marker for excision. T-DNAs are marked with kanamycin resistance (kan^r).

used for the isolation of two tagged mutants (25, 26) is demonstrative for the application of two element *Ac-Ds* systems. Briefly, plants homozygous for an *Ac* transposase source marked by β-glucuronidase (GUS^+) and kanamycin resistant (kan^r) markers were crossed with plants homozygous for a *Ds* element, expressing hygromycin phosphotransferase (HPT), inserted within a streptomycin resistance excision reporter gene located on the T-DNA (*Figure 2*). F_3 progenies were screened for mutations.

The *DRL1* (*d*eformed *r*oots and *l*eaves) locus tagged by *Ds* was identified in F_3 progeny derived from a population of 487 hygromycin and streptomycin resistant ($hyg^r/strep^r$) F_2 plants using the Δ*NaeIsAc* transposase source (25) (*Figure 2A*). This transposase source gives rise to about 5% germinal excision, and predominantly yields independent *Ds* inserts (18). Therefore all $hyg^r/strep^r$ F_2 plants are used to collect F_3 seed for mutant screening.

The *alb3* (albino) mutant (26) was isolated using a slightly different strategy (*Figure 2B*). As the CaMV 35S-*Ac* transposase source often yields just one independent *Ds* insertion per F_2 family (24), only one or a few $hyg^r/strep^r$ F_2 progeny was collected from each F_1 plant to produce F_3 progeny for mutant screening. To obtain lines with stable *Ds* insertions, GUS^- F_2 plants lacking the *Ac* transposase were chosen. The tagged *alb3* mutant was found in the progeny of 201 ($hyg^r/strep^r/GUS^-$) F_2 plants from 1678 F_1 individuals.

Unstable *alb3* mutants, resulting from somatic *Ds* excision, were found in progeny of GUS$^+$ (= *Ac* transposase$^+$) F_2 siblings of the *alb3* segregating line.

Marking *Ds* elements by a selectable marker is essential to obtain many plants that have inherited a *Ds* after excision. With the 35S-*Ac* transposase source, 90% of the hyg^r/strep^r F_2 plants contained a transposed *Ds* (24), whereas only 50% with the Δ*Nae*Is*Ac* transposase source (27). In the other hyg^r/strep^r plants, *Ds* excised from only one of two SPT::*Ds* T-DNA alleles, without subsequent reinsertion.

2.2.3 *En/Spm-I/dSpm* systems

Based on experiments with tobacco using a wild-type *En/Spm* element of maize (28) a one element *En/Spm* system was developed for *Arabidopsis*, that surprisingly gave a much higher transposition frequency as compared to tobacco or potato or to *Ac* in *Arabidopsis* (29). The germinal excision frequency averaged 7.5% and remained constant over a number of generations. Excisions of *En/Spm* occurred often independently from each other, with a high reinsertion frequency. A two element system in which a non-autonomous *dSpm* element was activated by two transposase genes, controlled by CaMV 35S promoters, has not been analysed in detail yet (29).

At the CPRO-DLO an '*in cis En-I* two element system' was developed (30, 31), which harbours both the *En* transposase source and a non-autonomous *I* element inserted in a HPT marked T-DNA. The expression of the two transposase genes is controlled by a CaMV 35S promoter fused to a truncated immobile *En* element. This system combines the advantages of the one and two element systems, having both continuous transposition and the ability to stabilize *I* elements. One transgenic plant, carrying two loci with multiple T-DNA inserts, was allowed to self-fertilize to generate different populations, carrying transposed *I* elements at many different positions. Instead of the germinal excision frequency, the frequency of independent transposition was assessed by DNA hybridization, as the fraction of unique novel inserts compared to all *I* element inserts in the progeny of a certain parental plant. This frequency varied from 5–10% in progeny of plants carrying over five *I* element copies, to almost 30% in progeny of single copy *I* element plants.

2.2.4 Tagging with *I* elements

In principle every generation of *I* element-carrying lines obtained after selfing can be screened for mutants. Outcrossing a mutant with wild-type Landsberg *erecta* (L*er*) for one or two generations will segregate out the transposase source, help to reduce the number of inserts, and yield stable mutants for the isolation of *I* element tagged genes.

The *En-I* was used for tagging and isolation of the *MS2* (male sterility) gene (30), and others including *CER1* (eceriferum) (32), *LFY* (leafy flowers) (M. V. Byzova, unpublished results), *AP1* (apetala), *ABI3* (abscisic acid-insensitive), *LEC1* (leafy cotyledon) (M. Koornneef, K. Léon-Kloosterziel, and A. J. M. Peters, unpublished results), *GL2*, *ANL2* (anthocyaninless) (51), *LAD* (late anther dehiscence) (M. A., unpublished results), *WIL* (wilting) (A. P., unpublished results), and *SAP* (sterile apetala). All of these mutants showed either somatic or germinal reversion in the presence of transposase.

2.3 Which system to use?

2.3.1 Random tagging

The use of *Ac-Ds* systems for random gene tagging is labour-intensive, because it requires an *in vitro* selection for excision and reinsertion. Assuming that 2–4% of all *Ds* inserts give rise to mutations with visible phenotypes (25, 26), screening 2500–5000 F_3 families carrying transposed *Ds* is expected to yield about 100 tagged mutants. This may seem a lot, however a saturation mutagenesis in *Arabidopsis* may require the generation and screening of over 100 000 inserts in F_3 as was estimated by Feldmann (4).

The *En-I* system seems to be more adapted for random tagging. With an independent transposition frequency of 10–30%, it is possible to generate a large population of different inserts in only a few generations. For example, starting with a single M_0 plant, which harbours a hemizygous *En* transposase source and ten different *I* element inserts, at least 10% of all *I* elements are expected to reside at a new location in the next generation. There will thus be on average 1000 different hemizygous inserts in 1000 M_1 plants. M_2 seeds of these 1000 M_1 plants can be harvested in bulk. Sowing 12 000 bulked seeds gives a 95% probability of recovering all new homozygous inserts (assuming equal seed set and viability). Theoretically, 100 000 different inserts can be obtained from 10 000 M_5 plants, when starting with 100 M_0 plants (as unrelated as possible) containing ten *I* elements at different positions. When M_0 plants hemizygous for the *En* transposase source are used, on average 37.5% of the M_2 plants will have stable homozygous *I* elements, reducing the chance of losing mutants that are not distinguishable due to a high excision frequency. Lines containing transposed *I* elements will become available through the seed stock centres.

2.3.2 Targeted tagging

An advantage of two element *Ac-Ds* systems lies especially in targeted gene tagging. Targeted tagging involves transposon mutagenesis of an already mapped locus. It has been employed in maize, using either linked or unlinked transposable elements (33, 34), and recently also in *Arabidopsis* (35). As with wild-type *Ac* in maize (36) most *Ds* elements transpose to positions genetically linked to their original genomic location (27). Choosing *Ds* elements closely linked to a target locus will increase the chance of tagging that locus when compared to using unlinked *Ds* inserts. Some of the *Ds* T-DNA inserts, available from the seed stock centres, are already genetically mapped (27, 37)

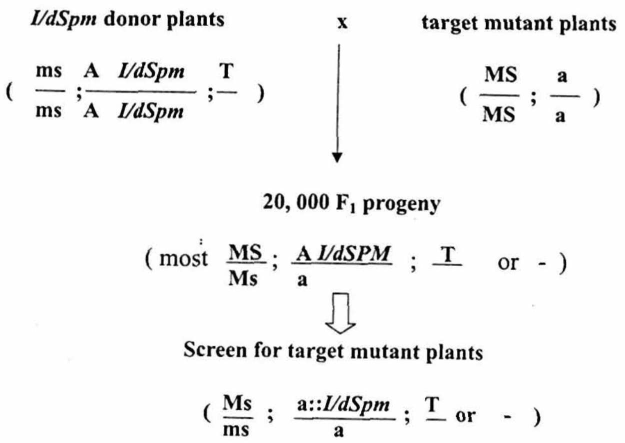

Figure 3. Strategy for targeted tagging of locus *A*, with a linked *I* element. *I* element donor plants are homozygous for a male sterile mutation (ms) and the *I* element linked to the *A* allele, and they are hemizygous for the *En* transposase (T). The target mutant plants are homozygous for the mutant *a* allele. Only F_1 plants with an *I* element insertion in the *A* allele (*a::I*) will show a mutant phenotype, that can be stable or unstable, depending on the presence or absence of the *En* transposase source (T or –).

and the appropriate ones can be chosen to start a targeted transposon tagging approach. When no *Ds* element is close (within 5–10 cM) to a target gene, such an insertion can be generated by mobilizing *Ds* elements from the closest mapped T-DNA carrying *Ds*. The new inserts can then be mapped relative to the T-DNA, using the T-DNA encoded excision marker and another marker encoded by *Ds* (27) as well as relative to other markers using RILs or YACs (as described in Chapters 3 and 8 respectively).

The high transposition frequency of the *En-I* also offers good possibilities for targeted gene tagging. Linked transposition of *I* elements was observed (31) although at a lower frequency than for *Ac-Ds*. An efficient strategy of targeted gene tagging is the use of male sterility to generate large F_1 populations (*Figure 3*). In this case a plant homozygous, for example, for an EMS-induced mutation at the target locus (A) is crossed with a nuclear male sterile plant (available from the seeds stock centres) (38), which is hemizygous for *En* transposase source and homozygous for an *I* element insert which is closely linked (< 5–10 cM) to the target A. In a large F_1 population heterozygous for the target A, homozygous mutants resulting from *I* element transposition into the target locus are screened or selected for. Using a transposon donor plant with one linked *I* element and a transposition frequency of 10%, a F_1 population of 20 000 plants is expected to carry 2000 new *I* element inserts, enriched for insertion near or at the target locus. As about 4000–5000 F_1 seeds can be obtained per male sterile plant (Aarts, unpublished results), five male sterile and five target mutant plants are sufficient to produce this required population. To obtain as many independent new insertion sites as possible, it is preferable to use more transposon donor plants (e.g. 20–50) with less crosses per plant.

2.4 Genetic and molecular analysis of a putatively transposon tagged mutant

A stepwise analysis (see *Protocol 1*) is required as a general strategy for cloning a transposon tagged gene. Initially an analysis of segregation of the elements and the mutation using revertants is necessary, prior to embarking on cloning (see *Protocol 1*). This protocol is applicable for the analysis of recessive mutations, which do not affect fertility. For rare dominant transposon-induced mutants, and infertile or inviable mutants (such as embryo lethals), these procedures will have to be adjusted.

For the analysis it is very important to use as few generations as possible. In each generation there is a chance of transposon excision, which in the worst case may generate a secondary transposon insertion closely linked to an empty, but mutated, target site. This may seriously complicate the genetic analysis. The observation of an unstable phenotype caused by excision, either as wild-type somatic sectors in a mutant background, or as wild-type germinal revertants in a mutant progeny, indicates a transposon-induced mutation.

Inverted PCR (IPCR) (39) is a reliable way to isolate DNA sequences flanking a transposon insert (*Protocol 2*). Plants with homozygous as well as hemizygous inserts can be used for IPCR. IPCR-derived flanking DNA fragments are used as probes for Southern DNA hybridization analysis of wild-type, mutant, and revertant plants to confirm the successful cloning of fragments from the tagged gene.

Protocol 1. Strategy for analysis of a putatively transposon tagged mutant

Equipment and reagents

- Liquid nitrogen
- −80 °C freezer
- Electrophoresis equipment and reagents (53)
- *Arabidopsis* plant material: wild-type line; line actively expressing transposase
- DNA isolation reagents
- Reagents for Southern hybridization (53)

Method

1. Harvest leaf material from the putatively transposon tagged mutant for DNA isolation (see *Protocol 2* or Chapters 3 and 8 for specific protocols). Cross (see Chapter 4 for conducting crosses in *Arabidopsis*) the mutant with wild-type (e.g. Landsberg *erecta*; cross 1) and with a transposase line if the mutant was not known to express transposase (cross 2).

Protocol 1. *Continued*

2. Follow segregation of the phenotype in the progeny and screen for presence of the transposase locus (e.g. by selecting for antibiotic resistance) (see Chapter 1 for growing *Arabidopsis* under selective conditions). If the mutant contains the transposase locus, screen 100–1000 progeny for wild-type looking revertants to test the stability of the mutant phenotype.
3. If the mutant does not contain a transposase locus, screen siblings to find a family expressing transposase. When found, screen the progeny as in step 2. Alternatively screen the F_2 from cross 2 for families with transposase. When found, screen progeny as in step 2. Only proceed when revertants are found.
4. Identify a transposon insertion co-segregating with the mutation. Preferably use a population without transposase (e.g. F_2 from cross 1), segregating 3:1 for wild-type:mutant. Alternatively, use a population with transposase, such as the F_2 from cross 2, or revertant and mutant progeny of the original mutant when the mutant contained the transposase locus. Perform a Southern blot analysis on about 50 plants for an *I* element tagged mutant (multiple inserts), or on about 10 plants (half mutant, half wild-type revertant) for a *Ds* tagged mutant (single or a few inserts). Load equal amounts of DNA per lane to distinguish between homozygous and hemizygous inserts. Make use of other populations if no co-segregating transposon can be identified.
5. Isolate genomic DNA flanking the co-segregating transposable element by IPCR (*Protocol 2*) using preferably DNA template from a plant lacking the transposase locus and carrying less than five copies of transposon inserts (especially for *I* element-induced mutants). If no plants with less than five inserts was found, use a backcross of the mutant with wild-type (F_1 cross 1 × wild-type) to reduce the transposon copy number.
6. Confirm the cloning of genomic DNA flanking the transposon by:
    (a) Hybridizing the IPCR probe to a Southern blot containing DNA from mutant and revertant plants, to reveal homozygous inserts in mutant and hemizygous or no inserts in revertant plants.
    (b) Analysis of the insertion site. Determine the sequence of DNA fragments carrying genomic DNA flanking the transposon insert using the IPCR fragments as template. Design PCR primers for the amplification of the insertional target site from wild-type DNA. PCR amplify the target site sequences from wild-type, revertant, and mutant alleles without inserts. Clone the PCR products and determine their DNA sequence. All revertants should have at least one allele with (near) wild-type DNA sequence. All mutants should have only alleles featuring frameshifts, aberrant termination, or amino acid exchanges.
7. Isolate genomic and cDNA clones from appropriate λ phage libraries using the IPCR products as probes, and determine their DNA sequence.

Homozygous mutant plants may not always contain a transposon insert in both mutant alleles. Occasionally, one of the inserts may transpose and leave an excision footprint behind, thus generating a stable mutant allele. When transposon insertions occur in coding regions, this is rather a rule than exception. Upon insertion, both *Ds* and *I* element generate a target site duplication of 8 bp and 3 bp respectively. Excision of these elements often deletes or duplicates a few base pairs leading to a frameshift or generating a stop codon. An adequate proof for cloning the correct gene can therefore be obtained by correlating the sequence of excision alleles with the plant phenotype. Revertant plants should have at least one allele for encoding a wild-type-like protein, whereas both alleles of a mutant should display an aberrant reading frame.

Protocol 2. Isolation of DNA probes flanking *I* element inserts in *Arabidopsis thaliana* by IPCR

Equipment and reagents

- Eppendorf tubes
- Liquid nitrogen
- Eppendorf-shaped grinders
- 14°C, 37°C, 65°C incubator
- Thermocycler
- DNA extraction buffer: 0.3 M NaCl, 50 mM Tris pH 7.5, 20 mM EDTA, 2% (w/v) sarkosyl, 0.5% (w/v) SDS, 5 M urea, 5% (v/v) phenol (equilibrated) (52); the first five ingredients are mixed as a 2 × stock solution, and urea and phenol are added before use
- Phenol (saturated):chloroform (1:1)
- 100%, 70% (v/v) ethanol
- Isopropanol
- Primers:
  1st PCR ILJ1 (120 ng/μl) GAA TTT AGG GAT CCA TTC ATA AGA GTG T
  IRJ1 (120 ng/μl) TTG TGT CGA CAT GGA GGC TTC CCA TCC GGG GA
  2nd PCR ILJ2 (120 ng/μl) ATT AAA AGC CTC GAG TTC ATC GGG A
  IRJ2 (120 ng/μl) AGG TAG TCG ACT GAT GTG CGC GC
  3rd PCR ITIR (105 ng/μl) GAC ACT CCT TAG ATC TTT TCT TGT AGT G
- TE: 10 mM Tris–HCl, 1 mM EDTA pH 8
- DNase-free RNase A (10 mg/ml stock)
- Enzymes: *Hin*fI, DNA polymerase I Klenow fragment, *Sal*I
- Spermidine (Sigma): prepare a 100 mM stock, store at −20°C
- 2.5 mM dNTPs (2.5 mM each dNTP)
- 0.3 M sodium acetate pH 5.5
- T4 DNA ligase
- 0.45 M NaCl
- *Taq* DNA polymerase
- 10 × PCR buffer
- 10 × TBE: 108 g Tris base, 55 g boric acid, 9.3 g Na_2EDTA per litre
- 1.2% (w/v) agarose (electrophoresis grade) in 0.5 × TBE

A. *Isolation of DNA from a single plant (52)[a] (especially for I elements)*

1. Harvest 100–150 mg of leaf or preferably inflorescence tissue per plant in an Eppendorf tube. Freeze in liquid N_2.

Protocol 2. *Continued*

2. Grind the tissue to a fine powder in the tube. Add 150 µl of DNA extraction buffer and grind once more. Add an additional 300 µl of extraction buffer and mix. Leave samples at room temperature until 18 or 24 samples are prepared.
3. Phenol:chloroform extract (450 µl) the samples. Precipitate the DNA with 0.7 vol. isopropanol. Keep tubes at room temperature for 5 min, then centrifuge for 5 min in an Eppendorf centrifuge at full speed. Wash the DNA pellet with 70% (v/v) ethanol and briefly dry them.
4. Dissolve the DNA pellets in 100 µl of TE containing 10 µg/ml RNase. DNA samples may be stored at 4 °C for a few months or at –20 °C.

B. *Preparation of DNA template for IPCR of I elements*

1. Digest 300 ng of DNA with 20 U *Hin*fI in 100 µl of 1 × *Hin*fI buffer containing 1 mM spermidine (3 h at 37 °C).
2. Add 1 µl of 2.5 mM dNTPs and 1 U of DNA polymerase I Klenow fragment. Incubate for 15 min at room temperature. Phenol:chloroform extract and precipitate the DNA with 0.1 vol. of 0.3 M NaAc and 1 vol. of isopropanol for 20 min at –20 °C (53).
3. Centrifuge the DNA for 20 min, wash the pellet in 70% (v/v) ethanol, and air dry. Resuspend in 99 µl of 1 × ligation buffer (53). Add 2.5 U of T4 DNA ligase and self-ligate the DNA fragments overnight at 14 °C.
4. Inactivate the ligase by heating the sample at 65 °C for 10 min. Use half of the DNA for step 5.
5. Add 25 µl of 0.45 M NaCl, 10 U of *Sal*I, and incubate for 3 h at 37 °C.
6. NaAc/isopropanol precipitate both non-treated and *Sal*I digested DNA samples (see step 4), wash the pellet with 70% (v/v) ethanol, and dry. Resuspend the DNA in 30 µl of sterile, distilled H_2O.

C. *Inverse PCR for I elements*

1. Transfer the DNA template into a PCR tube and add 4 µl of 10 × PCR buffer, 2 µl primer ILJ1, 2 µl primer IRJ1 (both at 120 ng/µl), and 2 µl dNTPs (2.5 mM each). Prepare 10 µl of 1 × PCR buffer with 2.5 U of *Taq* DNA polymerase.
2. PCR reaction:
   (a) 5 min at 95 °C (hot start).
   (b) Add 10 µl *Taq* DNA polymerase solution.
   (c) Set 25 cycles of PCR: 1 min 95 °C, 1 min 55 °C, 3 min 72 °C.
   (d) Elongate for 5 min at 72 °C.
3. Transfer 2 µl aliquot to a new PCR tube. Add 38 µl of 1 × PCR buffer, containing 2 µl primer ILJ2, 2 µl primer IRJ2 (both at 120 ng/µl), and 2 µl dNTPs (2.5 mM each).
4. Second PCR for 25 cycles using the conditions described in step 2.

D. *Cloning of IPCR fragments for I elements*

1. Size separate IPCR fragments on a 1.2% (w/v) TBE–agarose gel. Cut out the DNA bands from the gel, elute, and clone in an appropriate PCR cloning vector (53).
2. To obtain probes with very little *I* element sequence, use the (cloned) IPCR fragments for a third PCR with primer ITIR hybridizing to both terminal inverted repeats (TIR) of the *I* element.
3. Use 25 ng of linearized plasmid in a 50 µl PCR reaction (see part C), containing 2 µl of ITIR primer (at 105 ng/µl), but with annealing at 50 °C instead of 55 °C.
4. Clone PCR fragments as described in part D, step 1.

[a] Protocol adapted for *Arabidopsis* by Robert Whittier (personal communication) and used at CPRO-DLO for fast DNA analysis (Southern, PCR) of single plants.

Protocol 3. Isolation of DNA probes flanking *Ds* element inserts in *Arabidopsis thaliana* by IPCR

Equipment and reagents

- As *Protocol 2* (except for the following)
- Restriction enzymes: *Sau*3A or *Bst*Y1
- Primers (20 µM):

Ds 5′	1st PCR:	A3:	ATA CGA TAA CGG TCG GTA CGG G
		D74:	GGA TAT ACA AAA CGG TAA ACG GAA ACG
	2nd PCR:	D73:	TTT CCC ATC CTA CTT TCA TCC CTG
		E4:	CAA AAC GGT AAA CGG AAA CGG AAA CGG TAG
Ds 3′	1st PCR:	B39:	TTT CGT TTC CGT CCC GCA AGT TAA ATA
		B38:	GGA TAT ACC GGT AAC GAA AAC GAA CGG
	2nd PCR:	D71:	CCG TTA CCG ACC GTT TTC ATC CCT A
		D75:	ACG AAC GGG ATA AAT ACG GTA ATC

- 60 °C incubator
- Carrier yeast tRNA (10 mg/ml)

A. *Isolation of DNA from a single plant (52)*

1. Follow *Protocol 2*, part A.

B. *Preparation of DNA template for IPCR of Ds elements*

1. Digest 1 µg of DNA with 25 U *Sau*3A or *Bst*Y1 at 60 °C for 2–3 h in a 60 µl reaction volume. Remove 5 µl before adding the enzyme and compare with a 5 µl aliquot after the reaction on an agarose gel.
2. Ethanol precipitate the digested DNA. Centrifuge and wash with 70% (v/v) ethanol, air dry the pellet, and resuspend in 40 µl H_2O.

Protocol 3. *Continued*

3. Use 10 μl for self-ligation in 400 μl volume (2.5 U T4 DNA ligase, 14°C, overnight).
4. Add 2 μl of carrier tRNA (10 mg/ml) and 40 μl of 3 M NaAc pH 5.5. Phenol:chloroform extract and precipitate the DNA by adding 1 ml of cold 100% (v/v) ethanol to the supernatant, at −70°C for at least 30 min.
5. Centrifuge the DNA (10 min). Wash the pellets with 500 μl of 70% (v/v) ethanol, dry, and resuspend in 10 μl H_2O. Compare a 2 μl aliquot with a sample of digested, but unligated DNA by electrophoresis. Estimate DNA concentration by fluorimetry or A_{260}/A_{280}.

C. *IPCR for Ds elements*

1. Use 5 μl (50 ng) of DNA template in a 100 μl PCR reaction mix, containing 1 × PCR buffer, 5 μl of each first set primer (20 μM), 8 μl of 2.5 mM dNTP mix, and 2.5 U *Taq* DNA polymerase.
2. PCR reaction:
   (a) 5 min 94°C.
   (b) 35 cycles of 30 sec 94°C, 30 sec 55°C, and 3 min 72°C.
3. Use 5 μl for a further PCR reaction with a second set of primers, using conditions as in step 2.
4. Check 10 μl on a 2% (w/v) agarose minigel.

D. *Cloning of IPCR fragments for Ds elements*

1. Size separate IPCR fragments on a 1.2% (w/v) TBE–agarose gel. Cut out the DNA bands from the gel, elute, and clone in an appropriate PCR cloning vector (53).

2.5 Further applications of transposon tagging

2.5.1 Promoter or enhancer trapping with transposable elements

Less than 4% of *Ds* insertions were found to yield a visible mutant phenotype in *Arabidopsis* (25, 26). To benefit from the other 96%, transposons can be equipped with a reporter gene that is activated when the transposon is integrated in the vicinity of a transcriptional regulatory region, such as a promoter or enhancer sequence. In *Drosophila melanogaster*, P-elements containing a promoterless *lacZ* gene fused to the weak P-element transposase promoter are successfully used to detect transcribed genomic regions by transposon insertions (40). In *Nicotiana* and *Arabidopsis* this technique was first exploited using T-DNA insertions carrying promoterless reporter genes fused to the right T-DNA border (41–44, and discussed later).

Fedoroff and Smith (21) demonstrated the use of an *Ac-Ds* based promoter and enhancer trapping system in *Arabidopsis*. They constructed *Ds* elements which carried a promoterless *uidA* gene, linked either to a minimal promoter (the −46 to +6 region of the CaMV 35S core promoter) or directly to the 5′ terminus of *Ds*. Combination of a *Ds-uidA* element with a CaMV 35S-*Ac* transposase source in plants resulted in β-glucuronidase (GUS) expressing sectors after transposition. Recently a similar, but more advanced enhancer and gene trapping system has been described, in which plants are selected with *Ds* elements transposed to loci unlinked to the T-DNA donor locus (45, 46). Currently populations of plants containing independent transpositions are being built up to assay for GUS expression.

2.5.2 Insertion trapping by PCR

A novel technique, which may be widely used in *Arabidopsis*, exploits the abundance of transposons for identification of insertions in specific genes. Originally developed for *Drosophila melanogaster* (47) this PCR-based approach was taken to screen a library of P-elements in a fly population using two primers, one specific for terminal sequence of the P-element, the other derived from a target gene in which P-element insertions are desired. The resolution of PCR screening in *Drosophila* permitted the detection of one individual with the right insert in a population of 1000 flies. Analogously, with the aid of the Tc1 transposon in *Caenorhabditis elegans*, Zwaal *et al.* (48) have found 23 inserts for 16 different genes in a library of 960 worm cultures, pooled in a 10 × 10 × 10 three-dimensional matrix of 30 pools.

In plants this system has been applied to isolate mutants from *Petunia hybrida* using the dTph1 transposon (49, 76) and recently also in *Arabidopsis* using populations of plants containing T-DNA (50) or *En-I* transposable element inserts (82, 83). A prerequisite for this type of insertion–trapping, is a very high frequency of independent transpositions. For random inserts in *Arabidopsis*, only *En-I* systems seem to approach such high frequencies. Currently three such populations have been constructed. At the MPI für Züchtungsforschung, the AMAZE population of 8000 lines carrying 48 000 independent *En* insertions has been generated (83). At CPRO-DLO a population of 5000 plants carrying 45000 independent *I* element insertions has been prepared (Speulman *et al.*, in prep.) and at the Sainsbury lab of the John Innes Institute the SLAT population of about 48 000 lines has been grown in pools of 50, of which approximately 80% carry independent insertion events (84). Assuming a random distribution of inserts, the combined populations will on average contain an insert every 1 kb. With these transposon mutagenized populations it will not only be possible to obtain mutants for previously isolated genes by reverse genetics, but also to analyse mutations displaying very subtle phenotypes.

3. T-DNA tagging

3.1 The use of T-DNA as insertional mutagen

T-DNA tagging is based on a unique DNA transfer system of *Agrobacterium tumefaciens*, a soil borne plant pathogen, that causes grown galls mostly on dicotyledonous plants. *Agrobacteria* are capable of transferring a segment of their Ti or Ri (tumour-, or root-inducing) plasmids into plant cells. The transferred DNA (termed T-DNA) is flanked by 25 bp direct imperfect border repeats, and is stably integrated into the plant nuclear genome. Genes carried by the T-DNA are expressed in plants and encode functions for the synthesis of plant growth factors and specific metabolites (called opines), that can be used as sole carbon source by *Agrobacterium*. Deletion of oncogenes located between the 25 bp boundaries of the T-DNA does not affect the process of T-DNA transfer into plants, which is primarily regulated by Ti and Ri plasmid encoded virulence (*vir*) gene functions expressed in bacteria. Therefore, any foreign DNA can simply be transferred into plants by the help of T-DNA based vectors provided that the virulence gene functions are supplied *in cis* or *trans* in *Agrobacterium* (55). A wide variety of T-DNA based gene transfer vectors, as well as methods for transformation and regeneration of transgenic fertile plants, are now available for dicotyledonous, and few monocotyledonous species (see Chapter 6).

Molecular analysis of the T-DNA integration process revealed that the T-DNA is randomly integrated by illegitimate recombination into plant genomic loci that are potentially transcribed (43, 56, 57). It was thus predictable that, if T-DNA insertions occur frequently in genes, T-DNA induced insertional mutations causing visible phenotypes should likely arise when larger populations of T-DNA transformed plants are generated. The T-DNA was first used as a molecular tag to identify and isolate gene fusions in *Nicotiana* species and subsequently, when *Arabidopsis* became the model plant for molecular genetic research, a number of groups started generating large populations of T-DNA transformed *Arabidopsis* lines (4–6, 58).

3.2 Random tagging

As the T-DNA integration process is apparently not sequence-specific, T-DNA tagging is especially suited for random mutagenesis. In comparison to transposon tagging, an advantage of T-DNA tagging is that the T-DNA inserts are stable. Once integrated, the inserts remain at their original position, although recombination between multiple inserts may occur. By simply selfing transformants, large numbers of transformed seed can be obtained and distributed, so that many laboratories can screen the same T-DNA insertion library at different places over the world. With an efficient transformation method it is possible to generate a saturated population of T-DNA inserts. Every new transformant adds to the collection of T-DNA inserts, so that eventually the whole genome of *Arabidopsis* will be covered with insertions.

Feldmann (4) estimated that a population containing 105 000 randomly distributed T-DNA tags should be sufficient to saturate the *Arabidopsis* genome, to achieve a 95% probability of an average resolution of 2 kb between the inserts. Several laboratories using various transformation procedures contribute now to the approach of saturation mutagenesis. The available transformants can either be screened just for the segregation of mutants, or additionally for gene fusion insertions, depending on whether the T-DNA vector contains a reporter gene for detecting gene fusion insertions. In addition, T-DNA insertions in known genes (i.e. coding for expressed sequence tags, ESTs) can simply be identified by PCR aided screening of DNA pools from T-DNA insertional mutant lines (50, 76), and novel genes can efficiently be identified by random sequencing plant DNA fragments flanking the ends of T-DNA tags isolated by plasmid rescue or PCR amplification.

3.3 Available populations of T-DNA transformants

Table 2 summarizes the published populations of T-DNA transformants, four of which are available from the seed stock centres (38, 59) (see also Chapters 1 and 2). A very large population of transformants was obtained from the seed transformation experiments of Feldmann (4). Over 8000 transformants were generated in the Wassilewskija ecotype of *Arabidopsis*. On average 1.4 insertions are present in these transgenic lines, often represented by several T-DNA copies inserted into a single locus as direct or inverted repeats (4). Progeny from 4900 transformed lines are available from the seed stock centres as pools from 20 different transformed lines (38). Of the 8000 seed transformants screened, 15–26% segregated an offspring displaying visible

Table 2. Populations of T-DNA transformants

T-DNA selection	Gene fusion detection	Reporter gene	Bacterial *ori*	Population size	Ref.
Hygromycin	Transcription	*aph(3')II*	Yes	> 3000	62
Hygromycin	Translation	*aph(3')II*	Yes		62
Kanamycin	–	–	Yes	> 8000	4
Kanamycin	Transcription	*uidA*	No	171	41
Kanamycin	Translation	*uidA*	No	191	41
Kanamycin	Transcription	*uidA*	No	430	66
Kanamycin–hygromycin	–	–	No	> 1500	77
Kanamycin or basta	Transcription–translation	*uidA*	No	> 4000	65
Kanamycin	Transcription	*uidA*	No	12 000	78
Basta	35S readout	–	Yes	8550	79
Kanamycin	Transcription	GAL4-GFP	No	> 100	80

mutant phenotypes (4). The first gene that was isolated from this insertional mutant collection was *GL1*, a gene involved in trichome formation (60), but many more T-DNA tagged genes isolated from this population demonstrated the success of T-DNA tagging (5).

Another large population of over 3000 transformants was made in the Columbia ecotype by tissue culture transformation (61). A number of these transformed lines are also available from NASC (see Chapter 1) (38). From a small subset of 450 transformants, a *pale* mutant was identified, shown to be T-DNA tagged, and the corresponding *CH42* gene was isolated (62).

A problem thought to be associated with transformation by tissue culture methods is the generation of somaclonal mutants that have nothing to do with T-DNA insertion. However, it was found that non-tagged mutants were also frequently produced by other transformation methods, such as the seed transformation (4). In a screen of 1340 tissue culture transformed lines, 25.07% showed a mutant phenotype. Interestingly, the mutation frequency and the mutation spectrum reported are similar for the seed and tissue culture transformants (4, 6). The high mutation frequency seems very promising for gene tagging, but it has to be noted that many of the observed mutations did not co-segregate with a T-DNA insert. Castle *et al.* (63) performed an extensive characterization of 178 embryonic mutants derived from the seed transformants. They found that only 36% of the 115 mutants examined were actually tagged by T-DNA. Among tissue culture-derived transformants, Van Lijsebettens *et al.* (64) reported that only one out of seven mutants was T-DNA tagged. Koncz *et al.* (6) estimated that the proportion of T-DNA tagged mutants with an observable phenotype in their collection is 10–30%, stressing the importance of careful genetic linkage analysis before going into the process of gene cloning.

Recently, a very simple whole plant transformation procedure was published by Bechtold *et al.* (65) offering a practical possibility for high density gene tagging. Accordingly, these authors plan to generate a saturated T-DNA insertional mutant collection which they think will be reached by 50 000 to 100 000 independent insertions. A large number of these transformants are already available from the seed stock centres and many are expected to follow as an extended, valuable source of T-DNA tagged mutants.

3.4 Promoter/enhancer trapping

Many of the T-DNA transformed populations summarized in *Table 2* were generated using T-DNAs carrying a reporter gene for the detection of transcriptional or translational plant gene fusions. A major goal of using such promoter or enhancer trapping systems is either:

(a) To identify T-DNA insertions in coding regions using a selection or screening for the expression of translational fusions between plant genes and reporter genes.

(b) To detect T-DNA inserts in the vicinity of transcriptional regulatory elements that control gene expression spatially or temporally in response to developmental, hormonal, or environmental stimuli.

Gene fusions thus allow detection of gene mutations, without screening for a particular mutant phenotype, as well as permitting the analysis of gene expression in heterozygotes when insertional inactivation of a gene results in lethality. Koncz *et al.* (43), used a promoterless *aph(3')II* gene fused to the right T-DNA border in two variants, one with its own ATG start codon (and in-frame stop codons upstream), and another without ATG (and no in-frame stop codons upstream). Over 30% of the transformed plants tested expressed the APH(3')II kanamycin phosphotransferase reporter enzyme in different tissues. Similar experiments with the *uidA* reporter gene (for β-glucuronidase; GUS) resulted in the detection of 54% transcriptional and 1.6% translational fusions showing GUS activity in any tissue (44).

Transgenic lines expressing reporter gene fusions can be used to characterize promoters and their upstream regulatory sequences, as well as to isolate the genes corresponding to these sequences. To detect upstream regulatory sequences, Topping *et al.* (66) used a minimal TATA box promoter driven *uidA* gene fused to the T-DNA border. By assaying for *uidA* expression in siliques of 430 T-DNA transformants, they found 74 families displaying GUS activity. From one out of three transgenics showing embryo-specific GUS expression, they have isolated the genomic boundaries of the T-DNA insert and using these as probes, cloned the corresponding wild-type genomic and cDNA sequences. None of these lines with embryo-specific GUS expression resulted in aberrant phenotypes in homozygous offspring. Goddijn *et al.* (67) screened a similar T-DNA tagged population for down- and up-regulation of GUS expression in the syncytial cell during infection of *Arabidopsis* with nematodes. Insertional mutants found with the desired GUS expression will be used to identify regulatory sequences influenced by syncytial cell development. A large population of T-DNA tagged lines by Bechtold *et al.* (65) also contains a promoterless *uidA* gene for the detection of transcriptional and translational gene fusions, increasing its value for gene tagging experiments. Recently insertional mutagenesis and promoter trapping has been reviewed by Topping and Lindsey (68). For a practical overview of the generation of T-DNA induced reporter gene fusions in plants, the vectors to use for transformation and the cloning of regulatory sequences see Koncz *et al.* (69).

3.5 Analysis of T-DNA mutants and cloning a tagged gene

As described for the analysis of transposon-induced mutants, a stepwise protocol is given for the analysis of T-DNA induced mutations (*Protocol 4*) applicable for the collections reported in *Table 2*. The strategy is based on the assumption that the mutation is recessive, gives a clear phenotype, the homozygous mutant is fertile, and the T-DNA carries a dominant selection marker, such as antibiotic and/or herbicide resistance.

A problem encountered when screening a population of T-DNA transformants may be the occurrence of untagged mutations. It is therefore essential to rigorously confirm genetic linkage between the mutation and the T-DNA insert before attempting the isolation of the tagged locus. Because the T-DNA carries a dominant marker, rather large F_2/M_2 and F_3/M_3 populations have to be used to attempt the separation of untagged mutations from potentially closely linked T-DNA inserts. Finding no mutant without T-DNA among 1000 mutants means linkage within about 3.2 cM. *Protocol 4* describes the screening of mutants for presence of T-DNA. When working with a single T-DNA insert, alternatively the progeny of a wild-type plant (i.e. an M_2 family) carrying the T-DNA can be screened for mutants after selfing (i.e. this wild-type M_2 family is expected to be hemizygous for the T-DNA tagged locus). If within the T-DNA transformed progeny a wild-type family is found, not segregating mutant phenotype, the mutation is not T-DNA tagged.

When a different ecotype than the one used for transformation, is crossed with the mutant for making a segregating F_2 (provided the mutant phenotype is expressed in a different genetic background), ARMS (70) or CAPS (71) markers (see Chapter 3) can provide help to map the mutant to a chromosome arm.

Isolation of plant DNA fragments flanking the T-DNA by plasmid rescue (as described in *Protocol 5*) is only possible when the T-DNA contains an *E. coli* plasmid replication origin (*ori*). Otherwise the T-DNA insert junctions can be isolated by IPCR (see *Protocols 2* and *6*) using T-DNA-specific primers (69, 72). When the complementation of the mutant by transformation is complicated, for example when fertility is affected, different EMS or radiation-induced alleles can be sequenced and compared with the T-DNA locus (73).

Protocol 4. Analysis of putative T-DNA tagged mutants

Equipment and reagents

- Growth conditions and antibiotic selection for T-DNA
- Wild-type ecotype for crossing

Method

1. Make a segregating population by crossing the mutant with a wild-type ecotype. Select a F_1 plant containing the T-DNA marker, to produce around 4000 F_2 seeds.
2. Sow at least 100 F_2 plants on selective medium to estimate the number of T-DNA loci present in the mutant. If the segregation resistant: sensitive is significantly higher than 3:1, then cross several mutant families with the wild-type to produce several F_2 populations, at least one of which is segregating 3:1 for resistance:sensitivity. Self individual F_2 plants and screen F_3 families.
3. Sow 2000–4000 F_2 seeds of a properly segregating population in soil. Test for linkage by treating one or two leaves of 12 mutant plants with selective agent for T-DNA (determine the necessary amount first on wild-type plants).
4. When all 12 mutants are resistant, it is likely that the T-DNA is linked to the mutation (95% probability). Treat the rest of the mutant plants (about 500–1000) with the selective agent. Herbicide treatment is done by spraying, antibiotic treatment is safer to apply on leaves or, especially for hygromycin selection, by sowing progeny of mutants on antibiotic-containing medium (see Chapter 1). The mutation is not T-DNA tagged if a mutant is found without the T-DNA insert. Where possible score the mutant phenotype in Petri plates, alternatively germinate seeds on agar and transfer all mutants to antibiotic- or herbicide-containing media providing a selection for the T-DNA encoded dominant marker.
5. Isolate the T-DNA and flanking genomic DNA by plasmid rescue (*Protocol 5*) or IPCR (*Protocol 6*). Confirm the rescue and map transcript(s) using the flanking DNA fragments as probes for hybridizing Southern and Northern blots carrying DNA and RNA samples, respectively, prepared from wild-type, heterozygous, and homozygous F_2 or F_3 families.
6. Isolate genomic and cDNA clones from wild-type *Arabidopsis* libraries and perform their molecular characterization. At the same time accomplish the genetic or physical (YAC) linkage mapping of the mutation, using either classical or molecular methods as described in ref. 61 and elsewhere in this book.
7. Confirm the genetic linkage data by transformation of the T-DNA mutant with full-length genomic DNA and/or cDNA cloned in expression vectors, to demonstrate complementation. When available, use an EMS- or radiation-induced allele for transformation, to avoid silencing problems caused by multiple T-DNA inserts in the genome. Alternatively, characterize a number of mutant alleles (see text).

Protocol 5. Isolation of T-DNA flanking genomic DNA by plasmid rescue (69)

Equipment and reagents

- Reagents from *Protocol 2A*
- Equipment and reagents for Southern analysis (53)
- Bio-Rad Gene Pulser and cuvettes
- Agarose 0.8% (w/v) (electrophoresis grade) in 0.5 × TBE

Protocol 5. *Continued*

- 10 × TBE: 108 g Tris base, 55 g boric acid, 9.3 g Na$_2$EDTA per litre
- Phenol (saturated):chloroform (1:1)
- 3 M sodium acetate pH 5.5
- Isopropanol
- T4 DNA ligase
- 70% (v/v) ethanol
- LB medium (per litre): 10 g Bacto tryptone, 5 g Bacto yeast extract, 10 g NaCl
- Overnight *E. coli* culture (e.g. MC1061)
- 1 mM Hepes pH 7.0
- 1 mM Hepes pH 7.0, 10% glycerol
- SOC: 2% (w/v) Bacto tryptone, 0.5% (w/v) Bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl$_2$, 10 mM MgSO$_4$, 20 mM glucose
- Ampicillin, 100 mg/ml stock (filter sterilized)

A. *Preparation of DNA samples for electroporation*

1. Isolate DNA from mutant plants (see *Protocol 2A*, or ref. 74).
2. Determine by Southern analysis (53), which restriction enzyme is most suitable to use for plasmid rescue. The size of flanking DNA to be rescued should be optimally in the range of 1–4 kb.
3. Digest 5 μg of DNA (but as little as 100 ng can be used) with 25 U of the appropriate enzyme in 100 μl for 2 h. Check the digestion (5 μl) on an agarose gel.
4. Phenol:chloroform extract and precipitate the DNA with NaAc/isopropanol (10 min −20 °C), then centrifuge for 10 min at full speed.
5. Resuspend in sterile, distilled H$_2$O (20–50 μg/ml), add ligation buffer (53), and self-ligate overnight with 2.5 U T4 DNA ligase at 14 °C.
6. Phenol:chloroform extract, NaAc/isopropanol precipitate the DNA, wash twice with excess 70% (v/v) ethanol, dry, and resuspend in sterile, distilled H$_2$O at 10–100 μg/ml.

B. *Preparation of E. coli cells*

1. Inoculate 200 ml of LB with 2 ml of overnight *E. coli* (e.g. MC1061) culture. Grow for 2 h until OD$_{550}$ reaches 0.5, and centrifuge the cells (fixed-angle rotor) at 4 °C with 16 000 r.c.f. for 10 min.
2. Resuspend the cells in 100 ml of 1 mM Hepes pH 7.0 at 0 °C, recentrifuge, and resuspend in 50 ml of 1 mM Hepes. Recentrifuge and resuspend in 5 ml of 1 mM Hepes, 10% (v/v) glycerol.
3. Transfer the cells to Eppendorf tubes and pellet them at 4 °C (10 min, 2100 r.c.f.). Resuspend in 400 μl of 1 mM Hepes, 10% (v/v) glycerol.
4. Mix 40 μl of the cells with 10 μl of DNA in a pre-cooled cuvette for electroporation, e.g. a Bio-Rad Gene Pulser (25 μF, 2.5 kV, 200 A, and 4.8 msec). Immediately after electroporation add 1 ml of SOC (75) and incubate cells for 1 h with shaking at 37 °C.
5. Centrifuge cells briefly (20 sec) at full speed, resuspend in SOC, and plate aliquots on LB with ampicillin (100 μg/ml).

Protocol 6. Amplification and direct sequencing of T-DNA tagged plant DNA fragments by long-range IPCR (LA-IPCR)

Equipment and reagents

- Thermocycler
- Automated sequencer
- Equipment and reagents for CsCl banding (74–76)
- Qiagen DNA purification tip (optional)
- ABI Prism Dye Terminator Cycle Sequencing Kit (PE Applied Biosystems)
- Restriction enzyme that cleaves within T-DNA insert
- T4 DNA ligase
- Primers for nested PCR
- Elongase (Gibco BRL)
- 0.8% (w/v) agarose (electrophoresis grade) in 0.5 × TBE
- 10 × TBE: 108 g Tris base, 55 g boric acid, 9.3 g Na$_2$EDTA per litre
- Phenol (saturated):chloroform
- Isopropanol
- 0.3 M sodium acetate pH 6.0

Method

1. Purify high quality plant DNA by CsCl banding or on a miniscale with or without CTAB precipitation as described (see *Protocol 2*, and refs 74–76). Digest it with a restriction endonuclease which cleaves within the T-DNA insert.
2. Self-ligate the digested DNA (0.5 μg) as described in *Protocols 2, 3,* and *5*.
3. Optional: digest the ligated DNA with a restriction endonuclease which does not cleave within the T-DNA, but cleaves the plant DNA fragment flanking the left or right T-DNA end.
4. Design two sets of nested PCR primers:
   (a) One pair facing the T-DNA end (left or right).
   (b) Another pair facing the restriction endonuclease cleavage site within the T-DNA which was used in step 1.
5. Use half of the DNA in elongase PCR (BRL) or LA-PCR (Takara Shuzo Co.) for long-range amplification of large plant DNA fragments with a primer set facing the T-DNA end (left or right) and the endonuclease cleavage site within the T-DNA (step 1). Assemble the reaction mixes as recommended by the suppliers. Denature the template at 95 °C for 2 min, and perform 35 cycles of amplification (94 °C for 30 sec, 65 °C for 30 sec, 68 °C for 8 min), followed by elongation at 68 °C for 10 min.
6. Use one-tenth of the PCR mix to detect the product on an agarose gel, then gel isolate the PCR amplified fragments from the rest of the PCR mix.
7. If the yield is about 0.5 μg or more, purify each amplified DNA fragment with phenol:chloroform extraction and isopropanol precipitation (0.54 vol. isopropanol, 0.1 vol. 3 M Na acetate pH 6.0), or alternatively on a Qiagen tip. Use directly as a template for sequencing with an ABI

Protocol 6. *Continued*

Prism Dye Terminator Cycle Sequencing Kit and automatic DNA sequencer.

8. If the yield of PCR amplified plant DNA is low, dilute the PCR mix 500-fold and perform a repeated PCR reaction with the pair of nested primers located closer to the T-DNA end and the restriction endonuclease cleavage site used for IPCR in step 1.

3.6 Further applications of T-DNA tagging

At present T-DNA tagging is usually exploited to isolate genes from tagged mutants, by either screening for 'loss-of-function' type of mutations, or active 'promoter/enhancer trap' type T-DNA insertions, existing in the T-DNA transformed stock collections. None the less, T-DNA vectors designed for alternative applications are also available, and used for generation of new pools of transformants. One of these new applications, the activation tagging approach, is based on T-DNA constructs which carry transcriptional enhancer sequences linked to their termini. Upon integration, these transcriptional regulatory sequences may activate the expression of plant genes located in the vicinity of T-DNA inserts in a constitutive or developmentally regulated fashion (79). Many alternative designs using different promoters in combination with genes encoding transcription factors, suicide gene products, or different signalling factors are similarly expected to find application in studies of regulatory pathways and developmental processes in the future.

References

1. Bingham, P. M., Levis, R., and Rubin, G. M. (1981). *Cell*, **25**, 793.
2. Walbot, V. (1992). *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **43**, 49.
3. Haring, M. A., Rommens, C. M. T., Nijkamp, H. J. J., and Hille, J. (1991). *Plant Mol. Biol.*, **16**, 449.
4. Feldmann, K. J. (1991). *Plant J.*, **1**, 71.
5. Forsthoefel, N. R., Wu, Y., Schultz, B., Bennett, M. J., and Feldmann, K. A. (1992). *Aust. J. Plant Physiol.*, **19**, 353.
6. Koncz, C., Németh, K., Rédei, G. P., and Schell, J. (1992). *Plant Mol. Biol.*, **20**, 963.
7. Fedoroff, N. (1989). In *Mobile genetic elements* (ed. J. Shapiro), pp. 1–63. Academic Press, New York.
8. Peterson, P. A. (1970). *Theor. Appl. Genet.*, **40**, 367.
9. Pereira, A., Cuypers, H., Gierl, A., Schwarz-Sommer, Zs., and Saedler, H. (1986). *EMBO J.*, **5**, 835.
10. Masson, P., Surosky, R., Kingsbury, J., and Fedoroff, N. V. (1987). *Genetics*, **117**, 117.

11. Coen, E. S., Robbins, T. P., Almeida, J., Hudson, A., and Carpenter, R. (1989). In *Mobile DNA* (ed. D. E. Berg and M. M. Howe), pp. 413–36. Am. Soc. Microbiol., Washington DC.
12. Gerats, A. G. M., Huits, H., Vrijlandt, E., Maraña, C., Souer, E., and Beld, M. (1990). *Plant Cell*, **2**, 1121.
13. Tsay, Y.-F., Frank, M. J., Page, T., Dean, C., and Crawford, N. M. (1993). *Science*, **260**, 342.
14. Coupland, G., Plum, C., Chatterjee, S., Post, A., and Starlinger, P. (1989). *Proc. Natl. Acad. Sci. USA.*, **86**, 9385.
15. Baker, B., Coupland, G., Fedoroff, N., Starlinger, P., and Schell, J. (1987). *EMBO J.*, **6**, 1547.
16. Schmidt, R. and Willmitzer, L. (1989). *Mol. Gen. Genet.*, **220**, 17.
17. Dean, C., Sjodin, C., Page, T., Jones, J., and Lister, C. (1992). *Plant J.*, **2**, 79.
18. Bancroft, I., Bhatt, A. M., Sjodin, C., Scofield, S., Jones, J. D. G., and Dean, C. (1992). *Mol. Gen. Genet.*, **233**, 449.
19. Grevelding, C., Becker, D., Kunze, R., Von Menges, A., Fantes, V., Schell, J., *et al.* (1992). *Proc. Natl. Acad. Sci. USA*, **89**, 6085.
20. Swinburne, J., Balcells, L., Scofield, S., Jones, J. D. G., and Coupland, G. (1992). *Plant Cell*, **4**, 583.
21. Fedoroff, N. V. and Smith, D. L. (1993). *Plant J.*, **3**, 273.
22. Honma, M. A., Baker, B. J., and Waddel, C. S. (1993). *Proc. Natl. Acad. Sci. USA*, **90**, 6242.
23. Scofield, S. R., English, J. J., and Jones, J. D. G. (1993). *Cell*, **75**, 507.
24. Long, D., Swinburne, J., Martin, M., Wilson, K., Sundberg, E., Lee, K., *et al.* (1993). *Mol. Gen. Genet.*, **241**, 627.
25. Bancroft, I., Jones, J. D. G., and Dean, C. (1993). *Plant Cell*, **5**, 631.
26. Long, D., Martin, M., Sundberg, E., Swinburne, J., Puangsomlee, P., and Coupland, G. (1993). *Proc. Natl. Acad. Sci. USA*, **90**, 10370.
27. Bancroft, I. and Dean, C. (1993). *Genetics*, **134**, 1221.
28. Pereira, A. and Saedler, H. (1989). *EMBO J.*, **8**, 1315.
29. Cardon, G. H., Frey, M., Saedler, H., and Gierl, A. (1993). *Plant J.*, **3**, 773.
30. Aarts, M. G. M., Dirkse, W., Stiekema, W. J., and Pereira, A. (1993). *Nature*, **363**, 715.
31. Aarts, M. G. M., Corzaan, P., Stiekema, W. J., and Pereira, A. (1995). *Mol.Gen. Genet.*, **247**, 555.
32. Aarts, M. G. M., Keijzer, C. J., Stiekema, W. J., and Pereira, A. (1995). *Plant Cell*, **7**, 2115.
33. Nelson, O. E. and Klein, A. S. (1984). *Genetics*, **106**, 779.
34. Schmidt, R. J., Burr, F. A., and Burr, B. (1987). *Science*, **238**, 960.
35. James, D. W. Jr., Lim, E., Keller, J., Plooy, I., Ralston, E., and Dooner, H. K. (1995). *Plant Cell*, **7**, 309.
36. Greenblatt, I. M. (1984). *Genetics*, **108**, 471.
37. Coupland, G. (1992). In *Methods in Arabidopsis research* (ed. C. Koncz, N.-H. Chua, and J. Schell), pp. 290–309. World Scientific, Singapore.
38. Anderson, M. and Mulligan, B. (1994). *Seed list, 1994*. Nottingham Arabidopsis Stock Centre. Nottingham, UK (`http://nasc.nott.ac.uk`).
39. Ochman, H., Gerber, A. S., and Hartl, D. J. (1988). *Genetics*, **120**, 621.
40. Bellen, H. J., Wilson, C., and Gehring, W. J. (1990). *BioEssays*, **12**, 199.

41. André, D., Colau, D., Schell, J., Van Montagu, M., and Hernalsteens, J.-P. (1986). *Mol. Gen. Genet.*, **204**, 512.
42. Teeri, T. H., Herrera-Estralla, L., Depicker, A., Van Montagu, M., and Palva, T. (1986). *EMBO J.*, **5**, 1755.
43. Koncz, C., Martini, N., Mayerhofer, R., Koncz-Kalman, Zs., Körber, H., Redei, G. P., *et al.* (1989). *Proc. Natl. Acad. Sci. USA*, **86**, 8467.
44. Kertbundit, S., De Greve, H., Deboeck, F., Van Montagu, M., and Hernalsteens, J.-P. (1991). *Proc. Natl. Acad. Sci. USA*, **88**, 5212.
45. Sundaresan, V., Springer, P., Volpe, T., Haward, S., Jones, J. D. G., Dean, C., *et al.* (1995). *Genes Dev.*, **9**, 1797.
46. Springer, P. S., McCombie, W. R., Sundaresan, V., and Martienssen, R. A. (1995). *Science*, **268**, 877.
47. O'Hare, K. (1990). *Trends Genet.*, **6**, 202.
48. Zwaal, R., Broeks, A., van Meurs, J., Groenen, J. T. M., and Plasterk, R. H. A. (1993). *Proc. Natl. Acad. Sci. USA*, **90**, 7431.
49. Koes, R., Souer, E., van Houwelingen, A., Mur, L., Spelt, C., Quattrocchio, F., *et al.* (1995). *Proc. Natl. Acad. Sci. USA*, **92**, 8149.
50. McKinney, E. C., Ali, N., Traut, A., Feldmann, K. A., Belostotsky, D. A., McDowell, J. M., *et al.* (1995). *Plant J.*, **8**, 613.
51. Kubo H., Peeters, A. J. M., Aarts, M. G. M., Pereira, A., and Koornneef, M. (1999). *Plant Cell*, **11**, 1217.
52. Shure, M., Wessler, S., and Fedoroff, N. (1983). *Cell*, **35**, 225.
53. Sambrook, J., Fritsch, E. F., and Maniatis, T. (ed.) (1989). *Molecular cloning: a laboratory manual* (2nd edn). Cold Spring Harbor Laboratory Press, NY.
54. Altmann, T., Schmidt, R., and Willmitzer, L. (1992). *Theor. Appl. Genet.*, **84**, 371.
55. Zambryski, P., Tempe, J., and Schell, J. (1989). *Cell*, **56**, 193.
56. Mayerhofer, R., Koncz-Kalman, Z., Nawrath, C., Bakkeren, G., Crameri, A., Angelis, K., *et al.* (1991). *EMBO J.*, **10**, 607.
57. Gheysen, G., Villarroel, R., and Van Montagu, M. (1991). *Genes Dev.*, **5**, 287.
58. Valvekens, D., Van Montagu, M., and Van Lijsebettens, M. (1988). *Proc. Natl. Acad. Sci. USA*, **85**, 5536.
59. Anderson, M. A. (1995). *Weeds World*, **2**, 37.
60. Herman, P. L. and Marks, M. S. (1989). *Plant Cell*, **1**, 1051.
61. Koncz, C., Schell, J., and Rédei, G. P. (1992). In *Methods in Arabidopsis research* (ed. C. Koncz, N.-H. Chua, and J. Schell), pp. 224–73. World Scientific, Singapore.
62. Koncz, C., Mayerhofer, R., Koncz-Kalman, Zs., Nawrath, C., Reiss, B., Rédei, G. P., *et al.* (1990). *EMBO J.*, **9**, 1337.
63. Castle, L. A., Errampalli, D., Atherton, T. L., Franzmann, L. H., Yoon, E. S., and Meinke, D. W. (1993). *Mol. Gen. Genet.*, **241**, 504.
64. Van Lijsebettens, M., Vanderhaeghen, R., and Van Montagu, M. (1991). *Theor. Appl. Genet.*, **81**, 277.
65. Bechtold, N., Ellis, J., and Pelletier, G. (1993). *C. R. Acad. Sci. Paris, Sci. de la Vie/Life Sci.*, **316**, 1194.
66. Topping, J. F., Agyeman, F., Henricot, B., and Lindsey, K. (1994). *Plant J.*, **5**, 895.
67. Goddijn, O. J. M., Lindsey, K., van der Lee, F. M., Klap, J. C., and Sijmons, P. C. (1993). *Plant J.*, **4**, 863.
68. Topping, J. F. and Lindsey, K. (1995). *Transgenic Res.*, **4**, 291.
69. Koncz, C., Martini, N., Szabados, L., Hrouda, M., Bachmair, A., and Schell, J. (1994). In *Plant molecular biology manual* (ed. S. B. Gelvin and R. A. Schilperoort), pp. B2, 1–22. Kluwer Academic Publishers, Dordrecht.
70. Fabri, C. O. and Schäffner, A. R. (1994). *Plant J.*, **5**, 149.
71. Konieczny, A. and Ausuble, F. M. (1993). *Plant J.*, **4**, 403.
72. Lindsey, K., Wei, W., Clarke, M. C., McArdle, H. F., Rooke, L. M., and Topping, J. F. (1993). *Transgenic Res.*, **2**, 33.
73. Shevell, D. E., Leu, W.-M., Gillmore, C. S., Xia, G., Feldmann, K. A., and Chua, N.-H. (1994). *Cell*, **77**, 1051.
74. Dellaporta, S. L., Wood, J., and Hicks, J. B. (1983). *Plant. Mol. Biol. Rep.*, **1**, 19.
75. Zabarowsky, E. R. and Winberg, G. (1990). *Nucleic Acids Res.*, **18**, 5912.
76. Krysan, P. J., Young, J. C., Tax, F., and Sussman, M. R. (1996). *Proc. Natl. Acad. Sci. USA*, **93**, 8145.
77. Sangwan, R. S., Bourgeois, Y., and Sangwan-Norreel, B. S. (1991). *Mol. Gen. Genet.*, **230**, 475.
78. Jack, T. `http://www.dartmouth.edu/~tjack/`
79. Weigel, D. `http://www.salk.edu/LABS/pbio-w/index.html`
80. Haseloff, J. `http://www.plantsci.cam.ac.uk/Plantsci/Groups/JimHaseloff.html`
81. Byzova, M. V., Franken, J., Aarts, M. G. M., de Almeida-Engler, J., Engler, G., Mariani, C. *et al.* (1999). *Genes Dev.*, **13**, 1002.
82. Wisman, E., Hartmann, U., Sagasser, M., Baumann, E., Palme, K., Hahlbrock, K., Saedler, H. and Weisshaar, B. (1998). *Proc. Natl. Acad. Sci. USA*, **95**, 12432.
83. Baumann, E., Lewald, J., Saedler, H., Schulz, B. and Wisman, E. (1998). *Theor. Appl. Genet.*, **97**, 729.
84. Jones, J. D. G. `http://www.jic.bbsrc.ac.uk`