An adapter ligation-mediated PCR method for high-throughput mapping of T-DNA inserts in the *Arabidopsis* genome

Ronan C O'Malley[1], Jose M Alonso[2], Christopher J Kim[1], Thomas J Leisse[1] & Joseph R Ecker[1]

[1]Genomic Analysis Laboratory, The Salk Institute for Biological Studies, La Jolla, California 92037, USA. [2]Department of Genetics, Box 7614, North Carolina State University, Raleigh, North Carolina 27695, USA. Correspondence should be addressed to J.R.E. (ecker@salk.edu).

Published online 8 November; corrected online 13 December 2007 (details online); doi:10.1038/nprot.2007.425

***Agrobacterium* transfer DNA (T-DNA) is an effective plant mutagen that has been used to create sequence-indexed T-DNA insertion lines in *Arabidopsis thaliana* as a tool to study gene function. Creating T-DNA insertion lines requires a dependable method for locating the site of insertion in the genome. In this protocol, we describe an adapter ligation-mediated PCR method that we have used to screen a mutant library and identify over 150,000 T-DNA insertional mutants; the method can also be applied to map individual mutants. The procedure consists of three steps: a restriction enzyme-mediated ligation of an adapter to the genomic DNA; a PCR amplification of the T-DNA/genomic DNA junction with primers specific to the adapter and T-DNA; and sequencing of the T-DNA/genomic junction to enable mapping to the reference genome. In most cases, the sequenced genomic region extends to the T-DNA border, enabling the exact location of the insert to be identified. The entire process takes 2 weeks to complete.**

© 2007 Nature Publishing Group http://www.nature.com/natureprotocols

INTRODUCTION

Agrobacterium-mediated transformation in plants

Agrobacterium tumefaciens commonly carries a large plasmid, known as the Ti plasmid, which has the unusual ability to insert a piece of its DNA, the transfer DNA (T-DNA), into a plant's genome[1]. In the wild, the T-DNA carries a cassette of genes that convert the plant's tissue into a suitable environment for *Agrobacterium* colonization. In the laboratory, *Agrobacterium*-mediated T-DNA transfer has become a very useful tool for genetically transforming plants, including *Arabidopsis thaliana*[2,3].

One important application of T-DNA transfection is the creation of insertional mutants, plants in which a T-DNA has been inserted into a gene-coding region, thereby disrupting that gene's function. However, because the T-DNA insertion site in the genome cannot be experimentally controlled, it is impossible to target specific genes. As a result, it is necessary to create a large population of insertional mutants such that the genome can be saturated to the point that an insert is available in nearly every gene. Once a library of mutants has been generated and cataloged, researchers can directly order seed lines with a T-DNA insert in their gene of interest, greatly expediting gene function studies[4]. The primary advantage of T-DNA as a mutagen for creating large mutant libraries is that its sequence provides a target for a PCR primer, greatly simplifying the problem of mapping the precise location of the insertion site.

Indexing T-DNA insertion mutants by adapter ligation-mediated PCR

A critical step for creating an insertional mutant library is a method to identify the location of the T-DNA in a large number of individual lines. This is typically accomplished by sequencing the junction between the T-DNA insert and the unknown flanking genomic region, followed by mapping this junction sequence to the reference genome. To get good sequence information, it is critical to start with a sample that is heavily enriched for the T-DNA/genomic DNA (gDNA) junction region. This enrichment is best achieved by PCR. However, as the insertion site of the T-DNA is unknown, the PCR method employed needs to be able to capture the unknown genomic region flanking the T-DNA. As hundreds of thousands of individual seed lines must be screened to ensure good genomic coverage, the method must be very robust[4].

A global overview of our adapter ligation-mediated PCR strategy is provided in **Figure 1a**, whereas a sequence level closeup of the ligation and PCR steps are shown in **Figure 1b**. To begin with, gDNA containing a T-DNA insert (**Fig. 1a**, i) is digested with a restriction enzyme (**Fig. 1a**, ii), and adapters are ligated onto the overhanging ends of the gDNA (**Fig. 1a**, iii). The adapter (**Fig. 1b**) consists of a short strand (light blue) and a long strand (dark blue). When the two strands of the adapter are hybridized, the short strand contains a 5′ overhang that is complementary to the restriction site 'sticky end' on the gDNA and will ligate to the gDNA (**Fig. 1b**).

Many positions within the genome will be cut by the restriction enzymes and will subsequently ligate to the adapters (**Fig. 1a**, ii). Only one ligation site will produce the desired T-DNA–gDNA–adapter template (**Fig. 1a**, blue check mark), whereas the vast majority of products will contain genomic fragments with adapters ligated on both sides, adapter–gDNA–adapter (**Fig. 1a**, red X). If adapter-to-adapter PCR were possible, these targets would dominate the PCR and the T-DNA/gDNA junction would not be recovered. To prevent the unwanted adapter-to-adapter PCR, the ligated adapters are designed asymmetrically so that the long strand (**Fig. 1b**, dark blue) does not contain a binding site for the adapter primer, and so it is not capable of initiating a PCR. Instead, the long strand of the adapter contains an exact match to the adapter-specific primer, and only after its complementary strand has been synthesized will a binding site for the adapter primer be created. As a result, only the molecule containing the T-DNA-specific binding site will initially be copied, and only after this step will the adapter-binding site be created (**Fig. 1b**). In conclusion,

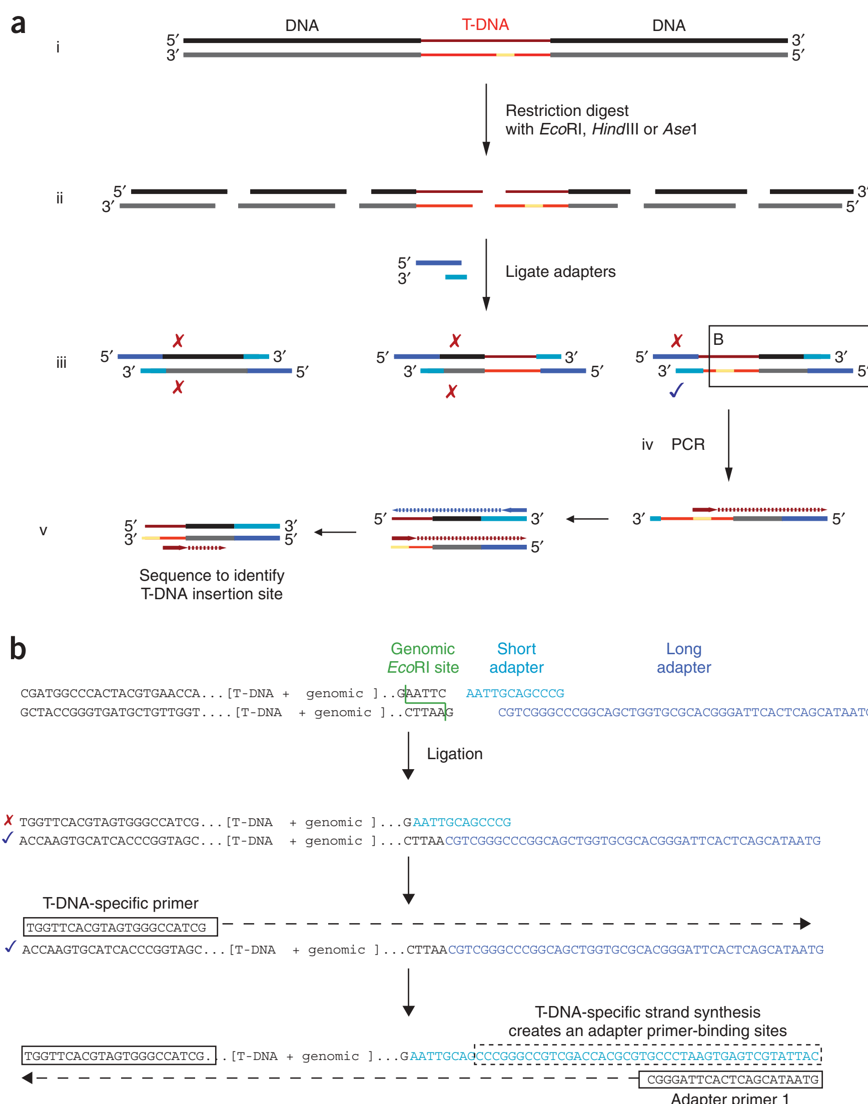

Figure 1 | Locating the T-DNA junction with adapter ligation-mediated PCR. (**a**) gDNA with a T-DNA insert (i) is digested with either *Eco*R1 and *Hind*III or *Ase*1 (ii). Adapters are ligated to the cut sites creating adapter-flanked templates (iii). The adapter sequence does not contain an adapter primer-binding site, but the longer arm of the adapter (dark blue) contains a 22 bp sequence that is an exact match to the adapter primer. Adapter-to-adapter amplification does not occur, because there are no adapter primer-binding sites (red X's). However, if there is a T-DNA present in the template (blue check mark), the T-DNA primer (red arrow) will bind to its corresponding site (yellow bar) and initiate synthesis of a complementary strand. The T-DNA primer PCR product will now contain an adapter primer-binding site derived from the complement of the longer arm of the adapter. The resulting moiety will have binding sites for both the T-DNA and adapter primers (blue arrow), so the target T-DNA/gDNA junction will be exclusively amplified (iv). After amplification, this moiety is sequenced with a nested T-DNA sequencing primer (red arrow) (v). (**b**) A sequence level view of the ligation of the Eco adapter and subsequent PCR with the AP1/LBa1 primer pair. To allow for the ligation of the adapter, the short strand of the adapter has been engineered to contain a 5′ overhang that is complementary to the overhang sequence of the *Eco*RI restriction digest. After ligation, the LBa1 T-DNA specific primer is the only primer that contains a binding site, and it will prime the synthesis of its complementary strand. However, when the long strand of the adapter is copied in this reaction, a binding site for the adapter-specific AP1 primer is created. This product now contains both the AP1- and LBa1-binding sites and the PCR will continue selectively on this product.

© 2007 Nature Publishing Group http://www.nature.com/natureprotocols

as complementary strand synthesis by the T-DNA primer is a prerequisite for the PCR to proceed, this strategy provides a means to selectively amplify the desired T-DNA–gDNA–adapter template (**Fig. 1a**, iv). Finally, after selective PCR amplification of the T-DNA/gDNA junction, the PCR product can be sequenced (**Fig. 1a**, v).

Alternative methods

An alternative approach for identifying T-DNA/gDNA junction sequence is thermal asymmetric interlaced PCR[5]. This method was used by Syngenta to identify the insertion sites in the Syngenta *Arabidopsis* insertion library T-DNA collection[6]. This approach employs T-DNA-specific primers and a pool of degenerate primers to randomly prime the complementary strand. To offset the high tendency for non-target priming by the degenerate set, multiple rounds of amplifications are performed with nested T-DNA primers and the same degenerate primer set. Thermal asymmetric interlaced PCR may have a higher tendency to produce false-positives owing to artifact amplification by PCR, although this method has proven to be effective and scalable, and provides an alternative to our ligated adapter method.

Applications of adapter ligation-mediated PCR

In addition to identifying T-DNA insertion sites, the adapter ligation-mediated PCR method described here has been adapted to identify the location of HIV virus integration in human gDNA to determine if there is any bias in the insertion site of this virus[7]. Over 524 unique insertion events were identified and preferential integration into genes was demonstrated. It would be expected that this method could provide a general, robust and scalable technique to locate random inserts of a known DNA element into a genome.

Experimental design

Choice of restriction enzyme. When selecting the restriction enzymes, there are two issues to consider. First, the restriction enzyme must not have a cut site in the region between the T-DNA primer-binding site and the T-DNA junction with the gDNA. If this region is cut by the restriction enzyme, the adapter will ligate within the T-DNA sequence and the desired T-DNA/gDNA junction will not be recovered.

The second issue to consider when selecting restriction enzymes is the distance between a cut site and the T-DNA insert. With a restriction enzyme that targets a 6 bp sequence, there will be on

TABLE 1 | Oligonucleotides required for this protocol.

Oligonucleotide name	Sequence (5′ to 3′)	Comments
Long strand of adapter 1	GTAATACGACTCACTTAGGGCACGCGTGGTCGACGGCC**CGGGCTGC**	Long strand adapter 1 is used in combination with short strand Eco and Hind adapters
Long strand of adapter 2	GTAATACGACTCACTATAGGGCACGCGTGGTCGACGGCC**CGGGCTGTGC**	Long strand adapter 2 is used in combination with the short strand of adapter Ase
Short strand of adapter Hind	5′-Phosphate-AGCT**GCAGCCCG-**amino C7-3′	5′ phosphorylated and 3′ C7 amino modification. HPLC purified
Short strand of adapter Eco	5′-Phosphate-AATT**GCAGCCCG-**amino C7-3′	5′ phosphorylated and 3′ C7 amino modification. HPLC purified
Short strand of adapter Ase	5′-Phosphate-TA**GCACAGCCCG-**amino C7-3′	5′ phosphorylated and 3′ C7 amino modification. HPLC purified
T-DNA left border a1 (Lba1)	TGGTTCACGTAGTGGGCCATCG	Primer for first PCR
Adapter primer 1 (AP1)	GTAATACGACTCACTATAGGGC	Primer for first PCR
T-DNA left border b1 (LBb1)	GCGTGGACCGCTTGCTGCAACT	Sequencing primer. Also the primer for nested PCR with AP2 for the Eco and Hind adapters
Adapter primer 2 (AP2)	TGGTCGACGGCCCGGGCTGC	Primer for second nested PCR for Eco and Hind adapters

Dotted underscore indicates the sequence complementary to the binding site for adapter primer 1. Solid underscore indicates the sequence complementary to the binding site for adapter primer 2. Bold text indicates the complementary regions between the long and short adapters that anneal to form the double-stranded adapters.

© 2007 Nature Publishing Group http://www.nature.com/natureprotocols

average a cut every 4,000 bp. In general, PCR amplicons experience a length-dependent decrease in amplification efficiency, so the more distant a restriction site is from a T-DNA insert, the more difficult it will be to identify that T-DNA insert. To maximize genomic coverage, the protocol can be repeated with different restriction enzymes. By using three different 6 bp restriction enzymes in separate reactions, a cut site should be present on average every 1,300 bp, providing good coverage for recovering T-DNA insert sites.

Adapter and primer design The adapters used in this protocol consist of two oligomers, a short strand and a long strand (**Table 1**). Both strands share a short complementary region to anneal the two strands (**Table 1**, bold text). After annealing, the short strand contains a 5′ overhang region that corresponds to the 'sticky ends' resulting from the digestion of the gDNA with the restriction enzyme. This overhang region is used for ligation of the adapter to the digested gDNA (**Fig. 1b**). The Hind, Eco and Ase adapters have overhang regions that correspond to their respective restriction enzymes (**Table 1**).

The Hind and Eco adapters are combined and ligated to their respective sites in a single reaction. The *Hin*dIII and *Eco*RI digestion and ligation steps are also performed simultaneously in this same reaction mixture. The Ase adapter, in contrast, is used by itself, and the digestion and ligation steps are performed separately. The reason for this arrangement is largely empirical.

The longer strand of the Hind and Eco adapters contains two latent primer-binding sites for the AP1 and AP2 primers (**Table 1**, dotted and solid underscore). Only after the sequence complementary to the long strand has been synthesized will the adapter primer sites be created (**Fig. 1b**). The short adapter has a 3′ C7 amino modification to prevent polymerase extension of the short adapter, as this would generate the latent primer-binding sites on all ligated molecules and not just those containing the T-DNA primer-binding site. For the Ase adapter, only one round of PCR is necessary, so only the AP1 primer is used. The exact sequence of the adapter primers is random, but tests should be run to ensure that the selected primers do not produce artifact bands.

Two different T-DNA primers, LBa1 and LBb1, are used in two rounds of PCR for the Hind/Eco adapter-ligated gDNA. The LBa1 and AP1 primer pair is used in the first PCR, and the LBb1 and AP2 pair is used in a second, nested PCR. The Ase adapter-ligated gDNA requires only one round of PCR with the LBa1 and AP1 primer pair. The LBb1 primer is used in all cases for the final sequencing of the T-DNA/gDNA junction. As there is some tendency for a small region of the T-DNA left border to be lost during incorporation, the T-DNA primer sites are set back from the left border of the T-DNA to ensure that they are present. However, the T-DNA-specific primers are also placed close to the left border of the T-DNA (>200 bp) to ensure that the amplicon is as small as possible for efficient amplification.

When designing the adapter primers or selecting the T-DNA primers, it is important to determine that the primers do not produce unwanted artifacts either by themselves or in primer pairs that will be used in the experiments. The best way to test this is to run PCR on a gDNA that has not been digested and ligated. If any of the primers alone or in the appropriate combinations produces a product, those primers should be replaced. A BLAST search of primer sequences against the target genome during the design stage may be helpful in identifying sequences with potential genomic-binding sites, but only a direct test can confirm the absence of false-positive bands.

MATERIALS

REAGENTS

- *A. thaliana* seeds, Col-0 (*Arabidopsis* Biological Resource Center)
- Bleach (Sigma-Aldrich)
- HCl (Sigma-Aldrich)
- Daconil 2787 (Ortho)
- MS salts (Invitrogen, cat. no. 11117-074)
- Gibberellic acid (1 mM stock solution) (Sigma cat. no. G-7645)
- gDNA wizard (Promega)
- Restriction enzymes *Hin*dIII, *Eco*RI, *Ase*I (20 U μl^{-1}) (NEB)
- T4 DNA ligase (20 U μl^{-1}) with buffer (NEB)
- ATP (10 mM) (NEB)
- *ExTaq* (5 U μl^{-1}) (TaKaRa), although any high-efficiency *Taq* should work.
- PCR-grade dNTP (10 mm each base) (NEB)
- Exonuclease I (10 U μl^{-1}) (NEB, cat. no. M0293S)
- ABI Big Dye version 3.1 (ABI Biotechnology, cat. no. 4336921)
- Ethyl alcohol (EtOH, 95% vol/vol) (Sigma-Aldrich)

- Sodium acetate (NaOAc) (Sigma-Aldrich)
- 5× sequencing buffer (400 mM Tris-HCl pH 9.0 and 10 mM $MgCl_2$)
- Liquid growth medium (see REAGENT SETUP)
- Oligonucleotides (Operon) (see **Table 1**). Prepare 5 μM stocks of all adapters and primers and store at −20 °C

EQUIPMENT

- 96 deep-well plate and lid (for plant growth) (Beckman)
- 96-well plate (for DNA wizard) (CLP, cat. no. 2407; or Beckman, cat. no. 267002)
- 96-well PCR plate (Sigma-Aldrich)
- Ziconaia/silica beads (Biospec Products)
- Thermocycler
- 37 °C incubator
- 5-liter hermetic chamber
- Seed spoon with 30 seed capacity (see EQUIPMENT SETUP and **Fig. 2**)
- Syringe and 0.22 μm syringe filter (Fischer)

REAGENT SETUP

Liquid growth medium To 100 ml of NanoPure H_2O, add 3 ml of Daconil 2787 (3.0% vol/vol) and 1 g MS salts (1.0% wt/vol). Autoclave and allow to cool. Add 60 μl of 1 mM gibberellin (0.6 μM) to the solution using a 0.22 μm syringe filter.

Preparation of 10× stocks of Hind and Eco adapters
Add 20 μl of 5 μM long strand adapter 1, and 20 μl of 5 μM short strand adapter (Hind or Eco) in 1,210 μl of 1 mM Tris, pH 8.3, in a 1.5 ml tube. Each adapter primer will be present at a final concentration of 80 nM. Vortex the tube and place it on a wet heat block at 96 °C for 2 min. Then turn off the heat and let the block cool to room temperature (20–24 °C); this should take about 10 min. This allows the long and short adapters to anneal. Before use, the success of the hybridization of the short and long strands of the adapter can be checked on a 20% (wt/vol) polyacrylamide gel. The hybridized adapters will run at a higher molecular weight than the short and long strands alone. Exact sizes may be difficult to determine, as the short and long adapter components will be single stranded when alone and the combination of the two will result in a partially double-stranded product. Adapter should be stored at −20 °C and can be kept for 6 months.

Preparation of a 10× stock of Ase adapter Resuspend 10 μl of 5 μM long strand adapter 2 and 20 μl of 5 μM short strand adapter (Ase) in 1,220 μl of 1 mM Tris, pH 8.3, in a 1.5 ml tube. The adapter primers will be present at a final concentration of 40 and 80 nM, respectively. Anneal, check and store the 10× stock as described above for Hind and Eco adapters.

EQUIPMENT SETUP

Creating a seed spoon A small depression is drilled into a metal bar (∼1 mm diameter). The hole size (∼0.5 mm diameter) is determined empirically on the basis of seed capacity (see **Fig. 2**).

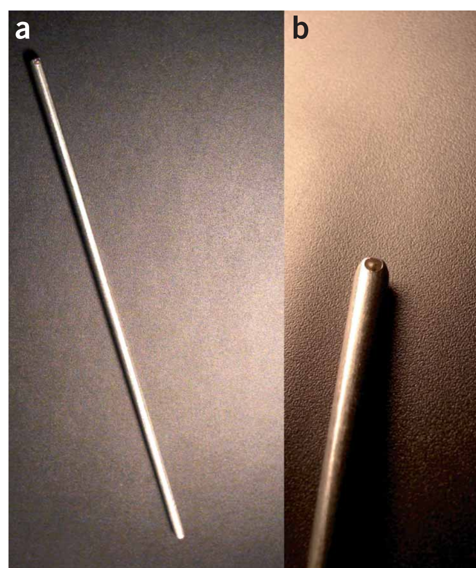

Figure 2 | Photograph of seed spoon. (**a**) ∼0.5 mm hole is drilled into an 18 cm metal bar with a 1 mm radius. (**b**) A close-up of the depression. A hole of this size can be used to reproducibly scoop out 20–40 *Arabidopsis* seeds.

PROCEDURE

Growing seedlings in a 96-well plate for DNA extraction

1| Transfer approximately 30 seeds from each independent T2 plant line to a separate well of a 96-well plate using a specially created seed spoon (see EQUIPMENT SETUP and **Fig. 2**).

2| Puncture the plate lid with a needle to allow air circulation in the wells. This is important during germination and seedling growth in Steps 9–11.

3| Place the seed plate (uncovered) with the lid in a 5-liter hermetic chamber in a fume hood.

4| Place a 50 ml beaker containing 25 ml of bleach into the chamber. Add 1 ml of concentrated HCl and immediately cover the chamber. The bleach and HCl react to release chlorine gas, which will sterilize the surface of the seeds but will not damage the embryo.
! CAUTION Chlorine gas is produced in this reaction. For safety, quickly seal the chamber and avoid inhaling any gas when the chamber is reopened in the fume hood.

5| After 30 min, remove the lid from the chamber and allow it to air out in the fume hood for 2 min; this allows the chlorine gas to escape from the fume hood.

6| Transfer the plate and lid to a laminar-flow sterile hood overnight to continue to air out.

7| The following morning, add 60 μl of liquid growth media.

8| Spin the plate at 1,500*g* for 15 s at room temperature to bring the growth media and seeds to the bottom of the wells. Cover with the lid.

9| Store the plate at 4 °C for 3 d in the dark to induce germination.

10| Transfer the plate to a growth chamber with constant light (4,500 lux) at 25 °C.

11| Let the seedlings grow for 6–8 more days, and then transfer the plate to the dark for another 3 d.
■ **PAUSE POINT** Seedling samples can be stored for up to 2 weeks at −20 °C before DNA extraction. DNA should be extracted from the sample soon after thawing, however, to avoid degradation.

© 2007 Nature Publishing Group http://www.nature.com/natureprotocols

© 2007 Nature Publishing Group http://www.nature.com/natureprotocols

gDNA extraction

12| Add one ball bearing to each well of the plate, either manually or with a 96-well ball-bearing dispensing tool. Add 150 µl of Promega wizard lysis buffer A to each well of the 96-well plate. Seal the plate tightly with two layers of aluminum foil tape. Place the 96-well plate into a Harbil 5G-HD paint shaker set to 200 Hz for 2.5 min. The rapid up-and-down motion shakes the ball bearings in each well and disrupts the plant tissue into the buffer. The leaf material should be completely shredded with only small, or no, pieces of leaf material visible to the eye, and the extraction buffer should be green.

13| Isolate DNA using the Promega wizard 96-well magnetic bead system for plants, according to the manufacturer's instructions.
■ **PAUSE POINT** Extracted DNA can be stored for up to 1 month at −20 °C.

Digestion of gDNA and ligation of adapters

14| Prepare the appropriate adapters (Hind and Eco or Ase adapters) as detailed in REAGENT SETUP.

15| If using the Hind and Eco adapters, which can be used together, follow option A; the gDNA can be digested and ligated to these adapters in a single step. If using the Ase adapter, follow option B. In this case, the gDNA must be digested before adapter ligation.

(A) Single-step digestion and simultaneous ligation of the Hind and Eco adapters

(i) Digest the gDNA and the Hind and Eco adapters and ligate these three components in a single step. First, add 3 µl (∼30 ng) of gDNA from each different sample from Step 13 to a separate well in a 96-well plate. Then add the components listed below to each well.

Component	Amount per well (µl)	Final
NanoPure H_2O	5.20	
10×NEB ligase buffer	1.00	1×
Hind adapter (from Step 14)	0.25	2 nM
Eco adapter (from Step 14)	0.25	2 nM
NEB *Hind*III (2 U rxn^{-1})	0.10	0.02 U
NEB *Eco*RI (2 U $rxn1^{-1}$)	0.10	0.02 U
NEB T4 DNA ligase (20 U rxn^{-1})	0.05	0.10 U
10 mM ATP	0.05	0.05 mM

(ii) Incubate overnight at room temperature.

(iii) To check the digestion of the gDNA, run a sample of undigested gDNA next to the digested sample on a 1% (wt/vol) gel. The digested sample should appear as a lower molecular weight smear averaging about 4,000 bp for the six-base targeting enzymes used in this protocol. The undigested sample will be observed as a smear at a much higher molecular weight.
■ **PAUSE POINT** Adapter-ligated DNA can be stored at −20 °C for 6 months.

(B) Two-step digestion and ligation (*Ase*I adapters)

(i) Add 3.0 µl (∼30 ng) of gDNA from each different genomic sample from Step 13 to a separate well in a 96-well plate.

(ii) Digest the gDNA by adding the following to each well of the 96-well plate:

Component	Amount per well (µl)	Final
Double distilled water	2.65	
10× NEB ligase buffer	0.65	1×
*Ase*I (2 U rxn^{-1})	0.20	0.06 U

(iii) Incubate for 4 h at 37 °C.

(iv) Ligate the Ase adapter to the digested DNA by adding the following to each well of the same 96-well plate:

Component	Amount per well (µl)	Final
10× Ase adapter stock	1.00	1×
Double distilled water	1.00	
10× NEB ligase buffer	1.00	1×
10 mM ATP	0.30	0.3 mM
T4 DNA ligase (20 U rxn^{-1})	0.20	1.0 U

(v) Incubate overnight at 20 °C. (vi) To check the digestion of the gDNA, run a sample of undigested gDNA next to the digested sample on a 1% (wt/vol) gel. The digested sample should appear as a lower molecular weight smear averaging about 4,000 bp for the six-base targeting enzymes used in this protocol. The undigested sample will be observed as a smear at a much higher molecular weight.
■ **PAUSE POINT** Samples can be stored for 6 months at −20 °C.
? TROUBLESHOOTING

© 2007 Nature Publishing Group http://www.nature.com/natureprotocols

PCR with the T-DNA and adapter primer pairs

16| Transfer 1 µl of each different adapter-ligated (*Ase*I or Hind/Eco) gDNA from Step 15 into the corresponding well of a clean 96-well plate.
▲ **CRITICAL STEP** A gDNA sample known to contain a T-DNA insert can provide a positive control for adapter-ligation PCR. A good negative control is a sample of gDNA that does not contain a T-DNA insert.

17| PCR-amplify the desired T-DNA/gDNA junction with the first T-DNA primer (LBa1) and adapter primer (AP1). Add 19 µl of the following mix to each well containing 1 µl of gDNA in the 96-well PCR plate.

Component	Amount per well (µl)	Final
Double distilled water	14.15	
10× *ExTaq* buffer	2.00	1×
10 mM dNTP	0.80	0.4 mM
LBa1 primer (5 µM)	1.00	250 nM
AP1 primer (5 µM)	1.00	250 nM
ExTaq (5U μl^{-1})	0.05	0.0125 U

18| Cover the plate with a rubber mat or foil tape and immediately load on a thermocycler paused at the initial 96 °C step cycle. This allows for a hot start to minimize initial mispriming at lower temperatures. Run the following PCR program:

Cycle number	Denature	Anneal and extend
1–10	96 °C for 0:20	72 °C for 2:20
11–25	96 °C for 0:20	67 °C for 2:20

■ **PAUSE POINT** The completed PCRs can be stored at −20 °C for 6 months.

19| Load 6.5 µl of PCR product on a 1.5% (wt/vol) agarose gel to visualize the products. Distinct bands should be present in most lanes when T-DNA/gDNA junctions have been recovered. Because there can be more than one insertion in a line, multiple but discrete bands may be present.
? TROUBLESHOOTING

20| If the *Ase*I adapter was used in the previous steps, proceed straight to Step 24 (the exonuclease reaction). If the Hind/Eco adapters were used in the previous steps, it is necessary to perform a second nested PCR to enrich for the target using the T-DNA primer LBb1 and Adapter primer AP2, as described in Steps 20–23. Transfer 0.2 µl of each first-round PCR product for the Hind/Eco reaction from Step 18 to the corresponding well of a clean 96-well PCR plate.

21| To each well of the plate, add the following:

Component	Amount per well (µl)	Final
Double distilled water	14.95	
10× HotMaster PCR buffer	2.00	1×
10 mM dNTP	0.80	0.4 mM
LBb1 primer (5 µM)	1.00	250 nM
AP2 primer (5 µM)	1.00	250 nM
ExTaq (5 U μl^{-1})	0.05	0.0125 U

22| Cover with rubber mat or foil tape and immediately load onto the PCR machine, paused at the initial 96 °C step. Run the PCR on a thermocycler using the following conditions:

Cycle number	Denature	Anneal	Extend
1–5	96 °C for 0:30	94 °C for 0:20	72 °C for 2:20
6–28	96 °C for 0:20	67 °C for 0:20	72 °C for 2:10

23| Load 2.5 µl of PCR product on a 1.5% (wt/vol) agarose gel to visualize the products. Distinct bands should be present in most lanes when T-DNA/gDNA junctions have been recovered. Because there can be more than one insertion in a line, multiple but discrete bands may be present.
■ **PAUSE POINT** The completed PCRs can be stored at −20 °C for 6 months.
? TROUBLESHOOTING

BOX 1 | CONCENTRATION AND PRECIPITATION OF PCR PRODUCTS

Optional concentration of PCR products before precipitation
If the exonuclease reaction is performed in a 384-well plate for the *Ase*I products, this additional concentration step is required before precipitation.

(i) Place the plate in the speed-vac at high heat for 1 h until the liquid is completely gone from the wells.
(ii) Add 10 µl of water to each reaction and let the DNA resuspend at room temperature for 1 h or overnight before precipitation. Overnight is preferable, as it ensures complete resuspension.

Precipitation of PCR products
(i) Transfer 10 µl of each PCR product to the corresponding well of a clean 96-well plate.
(ii) To each well, add 25 µl of 95% EtOH (vol/vol) and 1 µl of 3M NaOAc (pH. 4.7). The ethanol/NaOAc can be mixed as a stock in a 25:1 ratio.
(iii) Cover the plate and vortex.
(iv) Incubate at −80 °C for 15 min.
(v) Spin at a minimum of 3,000*g* for 25 min at 4 °C.
(vi) Remove supernatant by inverting plate on paper towel.
(vii) Wash pellets with 25 µl of 70% EtOH (vol/vol). Pellets are too small to be seen.
(viii) Spin for 10 min at 2,700*g* at room temperature.
(ix) Dry the DNA by inverting the plate on a paper towel and pulse the centrifuge to 500*g*.
(x) Add 10 µl of water to each reaction and let the DNA resuspend at room temperature for 1 h or overnight before precipitation. Overnight is preferable, as it ensures complete resuspension.

© 2007 Nature Publishing Group http://www.nature.com/natureprotocols

npg

Exonuclease treatment and precipitation of the PCR products

24| Exonuclease I is used to degrade the primers in the PCRs, as they can affect the quality of the sequencing reaction if not removed. For the Ase adapter PCR product, add the following to each well of the 20 µl PCR from Step 18. For the Hind/Eco adapter, the second PCR product from Step 22 is typically more concentrated and only 10 µl is required; transfer 10 µl of each reaction to the corresponding well of a clean 96-well plate and add 10 µl of double distilled water to each well before adding the components below:

Component	Amount per well (µl)	Final
Double distilled water	1.25	
10× Sigma PCR buffer	2.50	1×
10 *Exo*I (10 U μl^{-1})	1.25	0.5 U

25| Incubate at 37 °C for 1.5 h. To inactive the exonuclease, incubate at 80 °C for 20 min.
■ **PAUSE POINT** The exonuclease-treated product can be stored for 6 months at −20 °C or up to 2 years at −80 °C.

26| Precipitate and resuspend the PCR products in a 96-well plate as detailed in **Box 1**.
■ **PAUSE POINT** PCR products can be stored for one night at 4 °C and for 6 months at −20 °C before sequencing.

Sequencing the T-DNA-gDNA junction PCR products

27| Sequence with ABI Big Dye or any sequencing dye successfully used in your laboratory. If Big Dye is used, add the following components to each well of the precipitated PCR products from Step 26:

Component	Amount per well (µl)	Final
MilliQ H_2O water	7.00	
5× Sequencing buffer	2.00	1×
Big Dye	0.70	
LBb1 primer (5 µM)	0.30	150 nM

28| Cycle on Gene Amp 9700 with heated lid using the following program:

Cycle number	Denature	Anneal	Extend
1	95 °C for 1:00		
2–25	95 °C for 0:15	50 °C for 0:15	60 °C for 2:30

29| Precipitate and resuspend the sequenced products using the EtOH/NaOAc protocol described in **Box 1**.

30| Sequence the product on an ABI 3700 DNA sequencer, according to the manufacturer's instructions.

● **TIMING**

Germinate and grow seedlings in a 96-well plate for DNA extraction: Steps 1–8, 12–16 h (overnight); Step 9, 3 d; Steps 10 and 11, 8–12 d
Extract genomic DNA: Steps 12–13, 3 h
Prepare the Hind–Eco and Ase adapters, digest and ligate the adapters to genomic DNA: Steps 14 and 15, 12 h
PCR with the T-DNA and adapter primer pairs and electrophoresis: Steps 16–23, 6 h
Exonuclease treat, precipitate and sequence the T-DNA–gDNA junction products: Steps 24–30, 12 h

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 2**.

TABLE 2 | Troubleshooting table.

Step	Problem	Solution
15	Adapter ligation may fail when a new adapter stock is used. This will result in a failure of the PCR or sequencing of the T-DNA genomic junction. With a new adapter stock, it is wise to process a small number of samples to ensure that the adapter is working correctly	This may be a result of incorrect DNA concentration. Check the concentration by UV spectroscopy to confirm the concentration on the primer information sheet. Impure primers could also be the problem. It is useful to save a small amount of a successful adapter stock to compare to new stock by gel electrophoresis in case of problems. This stock can also be useful as a positive control if reagent troubleshooting becomes necessary
19 and 23	No PCR product in gel	PCR failure may be due to poor adapter ligation, although any component of the reaction may be at fault. Keep a set of active PCR reagents along with a positive control DNA sample in a −20 °C freezer. These can be stored for a year and retain full activity. If the adapter-ligated T-DNA template is not giving a product, substitute in the adapter-T-DNA template. Always run the control DNA to ensure that the stored PCR reagents are still active. If no product is produced from the adapter-ligated template, first check whether the restriction enzymes are working on a DNA sample, and if they are, then retest Step 15 with the stored, good adapter

© 2007 Nature Publishing Group http://www.nature.com/natureprotocols

ANTICIPATED RESULTS

We successfully used this protocol to map the location of T-DNA insertions in over 150,000 individual plants. In some cases, multiple T-DNA insertions are present in a single line. Because the LBb1 T-DNA primer is used as the sequencing primer, the T-DNA sequence of all the insertions will overlap and can be resolved. However, once the sequencing products reach the genomic sequence, each distinct product will produce its own sequencing signal, convoluting the sequencing information. However, once the sequencing has reached the point that only the longest T-DNA–gDNA product remains, the sequence will once again be resolvable, and it is possible to map this longest T-DNA product back to the genome. In this case, it will not be possible to locate the actual junction of the T-DNA/gDNA junction, as this sequence will be in the region of multiple overlapping reads. If the exact location is required, it is possible to design a sequencing primer within the newly identified gDNA to provide sequence going back into the T-DNA to confirm its presence.

ACKNOWLEDGMENTS This work was funded by the National Science Foundation (award numbers 0115103 and 0420126). We thank Cesar Barragan, Mary Galli, Dr Ryan Lister and Dr Brian Gregory for critical reading of this manuscript.

Published online at http://www.natureprotocols.com
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Chilton, M.D. *et al.* Stable incorporation of plasmid DNA into higher plant cells: the molecular basis of crown gall tumorigenesis. *Cell* **11**, 263–271 (1977).
2. Clough, S.J. & Bent, A.F. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743 (1998).
3. Hiei, Y., Ohta, S., Komari, T. & Kumashiro, T. 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 (1994).
4. Alonso, J.M. *et al.* Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* **301**, 653–665 (2003).
5. Liu, Y. *et al.* Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J.* **8**, 457–463 (1995).
6. Sessions, A. *et al.* A High-throughput *Arabidopsis* reverse genetics system. *Plant Cell* **14**, 2985–2994 (2002).
7. Schröder, A. *et al.* HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **110**, 521–529 (2002).