

Protocols

**A Simple Method for Identifying Plant/T-DNA
Junction Sequences Resulting from
Agrobacterium-mediated DNA Transformation**

*Yuanxiang Zhou, Ronald J. Newton and Jean H. Gould*¹

Department of Forest Science, Texas A& M University, College Station Texas
77843–2135

Key Words: Junction sequence, transformation, *Agrobacterium tumefaciens*

Abstract: Genomic integration of transferred T-DNA is traditionally analyzed by Southern hybridization; however, these analyses often do not provide sufficient information pertaining to the transformation event. Analysis of the junction sequences spanning the region between the T-DNA borders and plant genomic DNA, give a clear demonstration of genomic integration. The procedures available for border junction analysis can be problematic, therefore a simplified method was developed for plants transformed by *Agrobacterium tumefaciens* harboring the binary vector with pBI121 backbone.

Agrobacterium has been shown to have reliable efficiency in transferring genes into many different plant species. Inoculation of plant tissues with *Agrobacterium* results in induction of virulence genes, binding T-DNA with vir gene products and excision of the T-DNA at defined borders (RB: right border; LB: left border), producing single stranded T-DNA strands (T-strands) for transfer into plant cells. Once inside the plant cell, the T-strand appears to mimic a plant nuclear protein and is transferred to the cell nucleus by a highly conserved mechanism within the plant itself. In most case, this process eventually leads to integration of T-DNA into the plant chromosomes (Sheng & Citovsky, 1996)..

¹Author for correspondence.

Abbreviations: T-DNA, transfer DNA, RB, right border; LB, left border; PCR, polymerase chain reaction; IPCR, inverse PCR; pLP5, a plasmid resulting from inserting the Lp5 promoter of Loblolly pine in front of the 35S promoter of pBI121 from Clonotech.

Stable genomic integration of transferred genes is traditionally analyzed by Southern hybridization (Southern, 1975). However, analyses using only this technique sometimes do not provide sufficient information pertaining to the transformation event. Techniques such as plasmid rescue (Grant et al., 1990), inverse-PCR (IPCR) (Triglia et al., 1988; Does et al., 1991) and random primed PCR (Swensen, 1996, Trueba & Johnson, 1996) have been developed to identify the border junction sequences spanning the region between T-DNA and plant genomic DNA to clearly demonstrate true genomic integration.

Each of these methods present procedural problems. The plasmid rescue procedure has limitations in that knowledge of unique restriction sites in the transferred sequence is essential. In addition, the gene of interest is not often recovered because it does not encode an antibiotic-resistant protein or is lost due to recombination. In this case, the presence of a selectable marker does not faithfully mean that the transformation of the genes of interest has occurred. The IPCR procedure has multiple steps, and often the second digestion step cuts the fragment between the juncture and the first restriction site, abolishing the PCR amplification. The efficiency of self-ligation is also a limiting factor for the success of the technique.

The random-primed PCR method requires many different random primers, and generally, 20 to 40 primers are needed before one that amplifies the junction can be identified. Furthermore, random primed PCR often produce multiple bands, making Southern hybridization necessary to identify the band containing the junction sequences.

Here, we report a simple method to identify the junction sequences in plants transformed using by *Agrobacterium tumefaciens* using tobacco containing transformed with T-DNA from the binary vector pBI121 (Clontech). DNA isolated from the transformed plants is first digested by a *TaqI* restriction enzyme and ligated into an *AccI*-digested pUC18 plasmid; then the ligation mixture is subjected to PCR amplification performed by using one of the pUC universal primers and another primer which base pairs with the T-DNA sequence between the *TaqI* site and the right (RB) or left (LB) border. Finally, the PCR products are separated on an agarose gel, and the band is purified and sequenced. The procedure is outlined in Fig. 1.

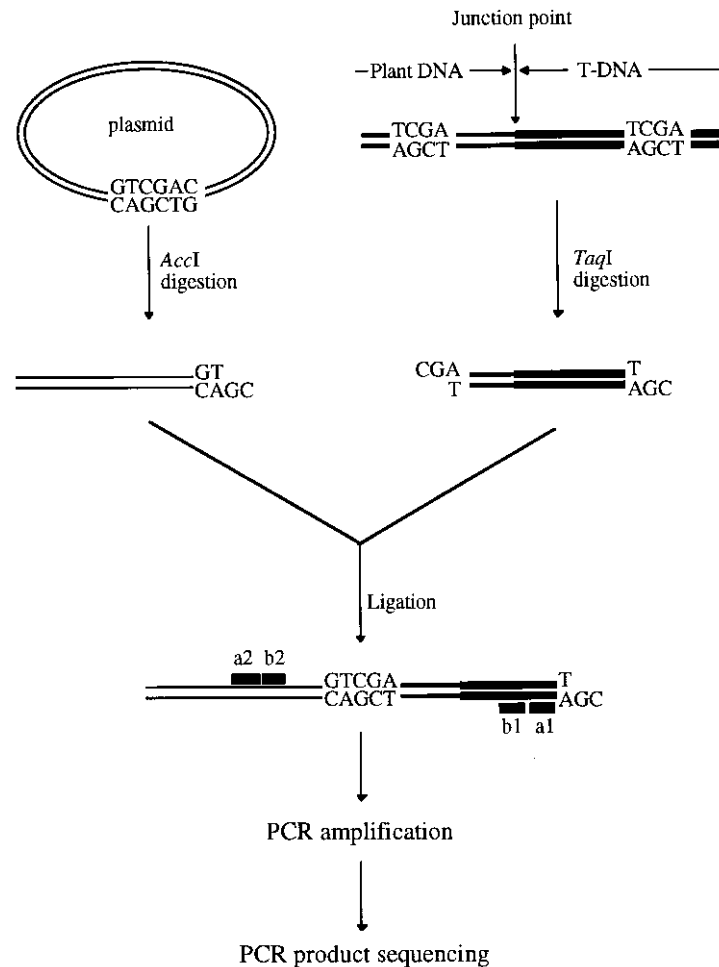


Fig 1. Ligation-mediated PCR for identifying junction sequences resulted from *Agrobacterium*-mediated DNA transformation. Plasmid pUC18 and genomic DNA from transgenic tobacco are digested with *AccI* (BRL) and *TaqI* (BRL), respectively. Both resulting restriction ends are compatible and can be easily ligated by T4 DNA ligase (BRL). Since cloning the *TaqI* fragment into the *AccI*-digested plasmid is not the purpose of this research, the molar ratio of both fragment ends are equal in the reaction mixture and the molar concentrations are higher than those recommended for cloning. The ligated fragments are PCR-amplified first with primers a1 and a2, and the resulting PCR products are then amplified with primers b1 and b2. The PCR products are separated on 1.2% agarose gel and visualized by ethidium bromide staining under UV light. DNA sequences are analyzed after purifying the PCR products from the gel.

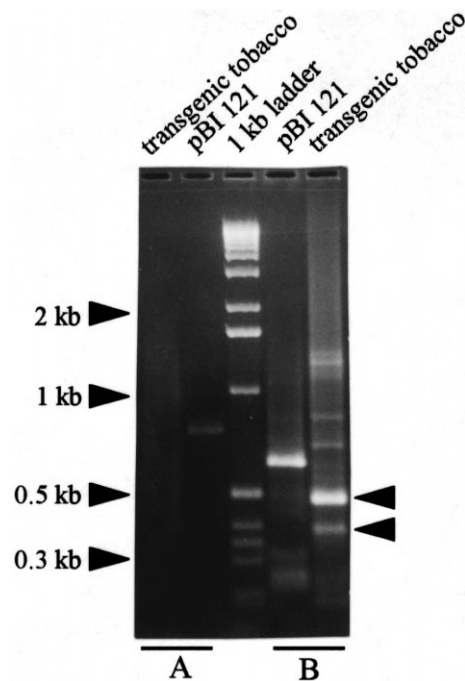


Fig 2. PCR amplification of ligated DNA samples. Ligation was performed with *AccI*-digested pUC18 and *TaqI*-digested pBI121, or DNA from transgenic tobacco. After ligation, PCR amplification is performed as described in the text. The PCR products are separated on a 1.2% agarose gel and visualized under UV light after staining. A. First PCR reaction; B Second (nested) PCR reaction. The two bands indicated by arrows to the right are PCR products containing junction sequences.

Materials and Methods

Plant materials: R1 population of tobacco transformed by *Agrobacterium* carrying the T-DNA in the binary vector pLP5 plasmid (Dias, 1995; Chang et al., 1996). When the plants are about 10 cm above ground, young leaves of 6–10 cm in length were harvested for DNA extraction.

Chemicals and equipment

Table-top centrifuge (Z229, National Centrifuge)
 Programmable Thermal Controller (PTC-100^{tiny}TM, MJ Research, Inc.)
 Waterbath (Model 183, Precision Scientific, Inc.)

Electrophoresis Power Supply (VMR 105, VMR Scientific; Model 3000Xi, Bio-RAD)
DNA fluorimeter (TKO 100, Hoefer Scientific Instruments)
X-ray film (Jersey Lab Supply)
Ethanol
Agarose (USB)
TE solution
10× TBE solution
QIAquick Gel Extraction Kit (Qiagen, Germany)
QIAquick PCR Purification Kit (Qiagen, Germany)
Phenol/chloroform/isoamyl alcohol (v/v/v, 25:24:1)
3 M sodium acetate (pH 5.2)
Mineral oil (Sigma)
*Taq*I restriction enzyme (BRL)
AccI restriction enzyme (BRL)
T4 DNA ligase (BRL)
Taq DNA polymerase (BRL)
dNTPs
pUC18 or pUC19 plasmid: extracted from the transformed *E. coli* strain DH-5 α
AmpliCycle™ Sequencing Kit (Perkin Elmer)

Protocol

Digestion of genomic DNA from transformed tobacco plants

One μ g genomic DNA is digested by at least 5 units of *Taq*I restriction enzyme in a 50 μ l mixture covered with mineral oil. The reaction is carried out at 65 °C overnight. After digestion, the mixture beneath the mineral oil is transferred to a 0.6 ml tube and extracted once with equal volumes of phenol/chloroform/isoamyl alcohol (25:24:1). The upper aqueous phase is transferred to another 0.6 tube, and DNA is precipitated after adding 1/10 volume of 3 M sodium acetate (pH5.2) and 2 volumes of cold ethanol. The pellet is washed with 70% ethanol and dissolved in 20 μ l TE solution. DNA concentration is determined with Fluorimeter TKO100.

Digestion of pUC18 plasmid by *AccI* restriction enzyme

One μg of pUC18 plasmid extracted from *E. coli* strain DH-5 α is digested by at least 5 units of *AccI* restriction enzyme at 37 °C overnight in 20 μl solution. After digestion, the mixture is separated on 0.8% agarose gel, and the band representing 2.6 kb, the size of pUC18 is cut out with a sterile blade. DNA is then extracted with a QIAquick Gel Extraction Kit, and the concentration is estimated as described above.

Ligation of *TaqI* digested tobacco genomic DNA with *AccI* digested pUC18

Ligation is performed at 16 °C for 4 h in a 20 μl solution which contains 150 fmol ends of *TaqI*-digested tobacco genomic DNA and 150 fmol ends of *AccI*-digested pUC18 fragment¹ (equal to 24.75 ng tobacco DNA and 132.9 ng pUC18, assuming the average length of *TaqI* fragments is 500 bp), 4 μl ligase buffer [(250 mM Tris-HCl (pH 7.6), 50 mM MgCl₂, 5 mM ATP, 5 mM DTT and 25% (w/v) poly ethylene glycol-8000)] and 0.5 unit T4 ligase (BRL). Five μl of the ligation mixture was used for the first round PCR. Therefore, 39.2 ng DNA (33.2 ng pUC18 and 6.2 ng tobacco DNA) was theoretically used in the first PCR amplification.

PCR amplification of the ligated DNA fragment

The PCR reaction is carried out in a 50 μl mixture which consists of 5 μl buffer [(200 mM Tris-HCl (pH 8.4), 500 mM KCl)] and 2 μl 50 mM MgCl₂, provided with the *Taq* polymerase kit (BRL); 1 μl of 10 mM each of dATP, dTTP, dCTP and dGTP; 25 pM each of the forward and reverse primers; 2.5 units of *Taq* polymerase. Two consecutive PCR reactions with different but nested sets of primers are performed for each ligated sample². The first is performed in a 50 μl mixture composed of 25 pM each of primer 5'-CGAATAGCCTCTCCACCCAAGCGGCCG-3' and primer 5'-TCGTATGTTGTGTGGAATTGGAGCGG-3' (a1 and a2, Fig. 1), 5 μl ligation solution as template, and other ingredients necessary for the PCR reaction as described above. The thermal cycler conditions are set at 94 °C for 45 sec for denaturation, 68 °C for 1 min for primer annealing, 72 °C for 1 min for elongation, for 40 cycles, except the initial cycle is 5 min for denaturation. This is followed by 72 °C for 10 min, and a 4 °C hold. After the PCR reaction, the products are purified by a QIAquick PCR Purification Kit to remove un-incorporated primers and nucleotides.

The second PCR reaction is carried out in 50 μ l solution using primers 5'-ATTGGATACCGAGGGG-3' and 5'-AACAGCTATGACCATG-3' (b1 and b2, Fig. 1), which are nested to the primers in the first PCR amplification. Two μ l aliquots of the first PCR product are used as template. The thermal conditions are the same as that for the first PCR reaction, except that the annealing temperature is 54 °C.

Electrophoresis and band purification

The products of the second PCR amplification are separated on 1.2% agarose gel and visualized with ethidium bromide under UV light. The bands are cut out and purified with the QIAquick Gel Extraction Kit³.

Sequencing PCR products

The purified PCR product is sequenced with a AmpliCycleO Sequencing Kit (Perkin Elmer). The sequencing primer is 5'-GATTGTCGTTTCCCGCC-3'. This primer anneals to fragments about 50 bp away from the beginning of the T-DNA right border. For each template, 1 μ l of α -³²P at the activity of 3000 Ci/mM is included in the reaction. The thermal conditions for sequencing are 1 min at 94 °C for denaturation, 1 min at 50 °C for annealing, 1 min at 72 °C for elongation, for 30 cycles, with the initial cycle being 94 °C for 5 min. After thermal reaction, 4 μ l stop solution are added, and the products are separated on 0.6% polyacrylamide gel, dried to a 3 MM sheet and exposed to X-ray film.

Notes

1. Theoretically, there should be one *TaqI* site for every 256 bp DNA fragment. After complete digestion of the tobacco genomic DNA with *TaqI*, restriction enzyme, we found a wide range of fragment sizes of the resulting restriction fragments following after agarose gel separation. The highest fragment density appeared between 400–1000 bp. In calculation of the fmol of the restriction fragments, an average size of 500 bp was used.
2. The necessity for nested PCR is due to the low copy of the desired fusion of vector and plant DNA during DNA ligation. Although the DNA concentration in terms of fmol ends in our ligation reaction was more than doubled in comparison with the manufacturer recommended concentration, the total genomic DNA used in a 20 μ l ligation solution was only a little bit more than 6 ng. A considerable portion of the ligation, should be self-ligation to form circular DNA and ligation between DNA of the same source (vector with vector, genomic with genomic). These ligation products will not result in PCR amplification. However, only ligation between vector and genomic fragment containing the junction sequences give PCR products, and therefore, the desired ligation may represent only a small fraction of the total

reaction product. The concentration of desired fragments from ligation are assumed to be insufficient to give satisfactory amplification in a single PCR reaction, and therefore, an additional PCR amplification, using nested primers, helps to amplify the fragments amplified in the first PCR reaction.

3. If multiple bands are produced, it is advisable to perform Southern hybridization to determine which bands should be sequenced. In case only one PCR band is visible, we recommend using the QIAquick PCR purification kit, or ExoI and phosphatase to remove remaining primers and unincorporated nucleotides.

Results and Discussion

The critical step in this method is to make the ends of the digested plasmid and plant DNA compatible. In addition, an appropriate distance between the restriction site and the LB or RB in the T-DNA is required so that probes can be made to correctly identify the PCR products in case multiple bands are produced.

Nested PCR primers are used because the first PCR amplification normally does not produce a visible band on agarose gel stained with ethidium bromide since the copy number of the fragments that can be amplified in the ligation mixture is too low. However, these fragments are amplified well above the limit of detection in the second PCR amplification. In our work with transformed tobacco, 2 to 4 bands are generally produced following the second PCR amplification. Gels are blotted and membrane hybridized with a probe made from the *Nos* promoter in the pBI121. Bands with homology to the probe are removed from the gel and sequenced. Sequences from putative transgenic plants are compared with the sequences adjacent to the RB and LB of the transforming binary vector. In transgenic plants, the DNA sequences immediately flanking the RB are not only different from those in the binary vector, but also not present in the T-DNA (sequencing data not shown). The results indicated stable integration of T-DNA in the tobacco genome.

In comparison with other procedures for investigating border junction sequences, the method described here is simple and time saving. Although we used pUC18 as a ligation cassette, virtually any kind of plasmid and vector can be used as long as a suitable restriction site is available.

Acknowledgements. This research was supported by Texas Higher Education Coordinating Board Advanced Technology Program grant 99990-0219 (RJ & JHG).

References

- Chang S., J.D. Puryear, E.A. Funkhouser, R.J. Newton and J. Cairney. 1996. Gene expression under water deficit in Loblolly pine (*Pinus taeda* L.). *Physiol. Plant.* 67:139–148.
- Dias M.A.D. 1995. Analysis of water deficit stress responsive cDNA clones and the characterization of the genomic clone of gene 'lp5' of Loblolly pine (*Pinus taeda* L.). Ph.D. Dissertation. Texas A& M University.
- Does M.P., B.M.M. Dekker, M.J.A. De Groot and R. Offringa. 1991. A quick method to estimate the T-DNA copy number in transgenic plants at an early stage after transformation, using inverse PCR. *Pl. Mol. Bio. Rep.* 17:151–153.
- Sheng J, Citovsky V. 1996. Agrobacterium-plant cell DNA transport: Have virulence proteins, will travel. *Plant Cell* 8:1699–1710.
- Southern E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503–517.
- Swensen J. 1996. PCR with random primers to obtain sequence from yeast artificial chromosome insert ends or plasmids. *BioTechniques* 20:486–491.
- Triglia T., M.G. Peterson and D.J. Kemp. 1988. A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences. *Nucleic Acids Res.* 16:8186.
- Trueba G.A. and R.C. Johnson. 1996. Random primed gene walking PCR: A simple procedure to retrieve nucleotide fragments adjacent to known DNA sequences. *BioTechniques* 21:20.