Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 8145–8150, July 1996
Plant Biology

Identification of transferred DNA insertions within *Arabidopsis* genes involved in signal transduction and ion transport

(reverse genetics/polymerase chain reaction/protein kinase)

PATRICK J. KRYSAN, JEFF C. YOUNG, FRANS TAX, AND MICHAEL R. SUSSMAN

Cell and Molecular Biology Program and the Department of Horticulture, University of Wisconsin, 1575 Linden Drive, Madison, WI 53706

Communicated by Hans Kende, Michigan State University, East Lansing, MI, April 17, 1996 (received for review February 16, 1996)

ABSTRACT **The transferred DNA (T-DNA) of *Agrobacterium tumefaciens* serves as an insertional mutagen once integrated into a host plant's genome. As a means of facilitating reverse genetic analysis in *Arabidopsis thaliana*, we have developed a method that allows one to search for plants carrying T-DNA insertions within any sequenced *Arabidopsis* gene. Using PCR, we screened a collection of 9100 independent T-DNA-transformed *Arabidopsis* lines and found 17 T-DNA insertions within the 63 genes analyzed. The genes surveyed include members of various gene families involved in signal transduction and ion transport. As an example, data are shown for a T-DNA insertion that was found within CPK-9, a member of the gene family encoding calmodulin-domain protein kinases.**

Over the past several years, plant biologists have amassed DNA sequence information for thousands of different genes and gene families from many species, including *Arabidopsis thaliana*. Missing from the plant biologist's tool box, however, has been a method for easily generating a plant that carries a "knockout" mutant of a given sequenced gene. This process of "reverse genetics" would greatly assist in efforts to determine the function of a particular gene *in vivo*. In this paper, we present a method for applying reverse genetics to *Arabidopsis thaliana* and report the identification of 17 transferred DNA (T-DNA) insertions within genes involved in signal transduction (protein kinases, calmodulins, 14–3-3 related proteins, and a phospholipase C) and ion transport (H^+ ATPases and a K^+ channel).

Our reverse genetic strategy involves starting with a large collection of T-DNA-transformed *Arabidopsis* lines. T-DNA is a segment of DNA from *Agrobacterium tumefaciens* that becomes integrated into a plant's chromosomes at seemingly nonspecific locations upon transformation by the bacteria (1). Once integrated, the T-DNA serves as an effective insertional mutagen with a known DNA sequence for which PCR primers can be designed (2). One can therefore detect the presence of a T-DNA element within a particular plant gene by performing PCR reactions using one primer targeted to the plant gene and a second primer targeted to the T-DNA. By screening a large enough number of T-DNA-transformed *Arabidopsis* lines, it should be possible to find a plant that carries a T-DNA fortuitously inserted within any gene of interest. Since gene disruptions are generally recessive, heterozygous knockouts of most *Arabidopsis* genes should result in healthy plants that are adequately represented in the T-DNA-transformed population.

Recently, McKinney *et al.* (3) published a strategy in which degenerate PCR primers were used to screen for T-DNA insertions within members of the actin gene family. We have developed a more general procedure based on gene-specific, nondegenerate PCR primers. In addition, we directly test the generality and sensitivity of our method by searching for T-DNA inserts within 63 different genes. The results reported in this paper demonstrate that our method is capable of detecting most, if not all, of the T-DNA inserts expected to be present in the available T-DNA-transformed population. The procedure as described in this paper should therefore be immediately useful to all researchers interested in searching for a T-DNA insert in any sequenced *Arabidopsis* gene.

MATERIALS AND METHODS

Plants. Of the 64 T-DNA transformed *Arabidopsis thaliana* pools of 100 available from the Arabidopsis Biological Resource Center (ABRC), 63 pools were used in this study along with 27 additional pools of 100 provided by E.I. DuPont deNemours. ABRC pool 2611 was omitted from this study due to insurmountable microbial contamination. Approximately 6000 seeds from each pool of 100 were surface sterilized and placed in 250-ml flasks with 75 ml of Gamborg's B5 medium, 2% (wt/vol) glucose, and 0.5% (wt/vol) Mes (pH 5.7). These flasks were placed on a shaker at *ca.* 125 rpm at 4°C for 4–6 days and then moved to a shaker at *ca.* 125 rpm at 22°C under continuous light for 7–14 days. Seedlings from each flask were then collected, blotted to remove excess liquid, and ground to a powder in liquid nitrogen by use of a chilled mortar and pestle. The pools of 20 were handled in a similar fashion.

Genomic DNA Preparation. Extraction buffer [500 μl; 0.2 M Tris·HCl, pH 9.0/0.4 M LiCl/25 mM EDTA/1% (wt/vol) SDS] and 500 μl of phenol/chloroform/isoamyl alcohol (24:24:1) were placed in a 1.9-ml microfuge tube. Frozen ground plant material (200 mg) was added, and the tube was mixed for 10 s with use of a vortex. After incubation on ice for 5 min, the tube was spun in a microfuge at 4°C for 5 min. A phenol extraction was then performed, and the DNA was precipitated with an equal volume of isopropanol and pelleted by centrifugation. This DNA pellet was resuspended in 500 μl of 10 mM Tris·HCl, pH 8/100 mM NaCl/1 mM EDTA plus 40 μg/ml RNase A. Another phenol extraction was performed, and the DNA was again precipitated. The DNA pellet was resuspended in a final volume of 500 μl of 10 mM Tris·HCl, pH 8/1 mM EDTA. This procedure typically yields 10 μg of DNA per preparation.

When DNA was isolated from individual plants, a modified version of the above procedure was used. Fresh leaf tissue was placed directly into a tube containing 30 μl of extraction buffer and ground using a drill mounted plastic grinding bit. Extraction buffer (470 μl) was then added to the tube along with 500 μl of phenol/chloroform/isoamyl alcohol. The remainder of the procedure was as described above.

PCR. The following PCR primers were used: T-DNA left border (L), gatgcactcgaaatcagccaattttagac (3); T-DNA right

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: T-DNA, transferred DNA; ABRC, Arabidopsis Biological Resource Center; CPK-9, calmodulin-domain protein kinase isoform number 9; L, left border primer; R, right border primer; 5′ and 3′, primers near the 5′ end and the 3′ end of the CPK-9 gene, respectively.

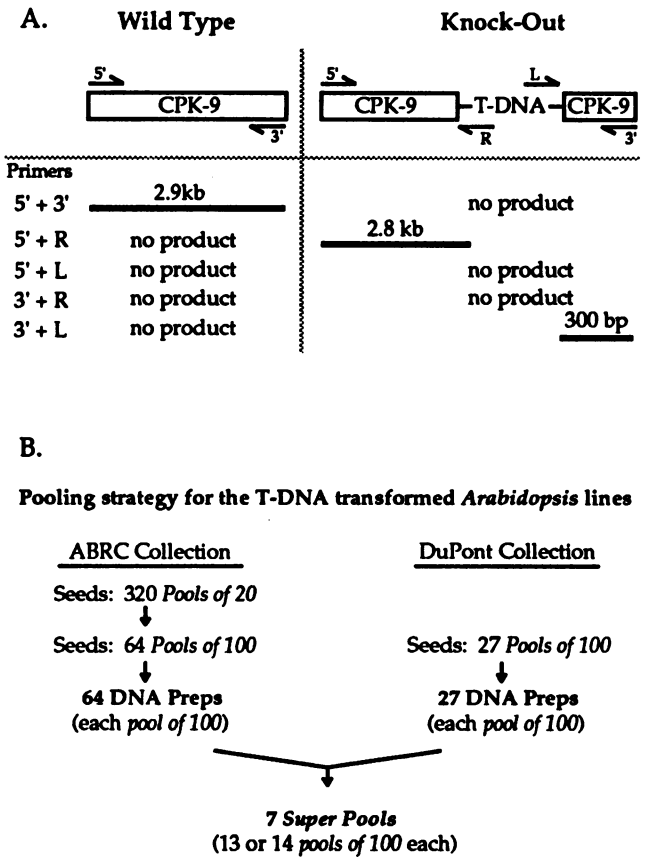


FIG. 1. Strategy for screening 9100 T-DNA-transformed *Arabidopsis* lines. (*A*) A schematic of the PCR products generated from CPK-9 with various oligodeoxynucleotide primer combinations is shown. Small arrows denote the direction and position of the PCR primers. L, T-DNA left border primer; R, T-DNA right border primer. T-DNA and the CPK-9 gene are not drawn to scale. (*B*) Organization of the T-DNA-transformed lines as pools of 20 and pools of 100 is shown together with the organization of DNA preparations into seven super pools.

border (R), tccttcaatcgttgcggttctgtcagttc; *c*almodulin-domain *p*rotein *k*inase isoform number 9 (CPK-9) 5′, aagttccgatctttc-ctctctgggtcgtc; CPK-9 3′, gaatcacattaaaagtggaggaactgagg. PCR was performed using PanVera X-Taq polymerase and the supplied X-Taq buffer. This enzyme was chosen because it allowed us to amplify large products, which meant that only two gene-specific primers were needed to analyze an entire gene. Each 50-μl reaction contained 40 ng of DNA, 12 pmol of each primer, 3 units of X-Taq enzyme, and 50 μl of mineral oil. In the hot start procedure that was used, 40 μl of the PCR reaction, including all of the components except the X-Taq, were heated to 96°C for 5 min. Ten microliters of 1× X-Taq buffer plus 3 units of X-Taq polymerase were then added. The following PCR program was used: 96°C for 5 min and then pause at 96°C for as long as it takes to add enzyme to each tube; 41 cycles of 94°C for 30 s, 65°C for 1 min, 72°C for 2.5 min; 72°C for 5 min; 4°C hold. A Perkin–Elmer model 2400 or 9600 thermal cycler was used. PCR products were cycle sequenced and run on a Perkin–Elmer/ABI Model 377 automated DNA sequencer.

PCR reactions were separated on 0.8% to 1.0% (wt/vol) agarose gels, stained with ethidium bromide, and photographed. Hybridization of the Southern blot was carried out under stringent conditions with radioactively labeled PCR product from a reaction that used the CPK-9 primers 5′ and 3′ to amplify a fragment of the wild-type CPK-9 gene.

RESULTS

PCR Strategy. The method that we describe in this paper should be generally applicable to any *Arabidopsis* gene for which some DNA sequence is known. As the rationale and implementation of the technique are best described by example, we will present a detailed account of how the process was used to find a T-DNA insertion within the CPK-9 gene (4). The basic PCR strategy is outlined in Fig. 1*A*. First, two PCR primers (5′ and 3′) that were specific for the CPK-9 gene were designed. Amplification of wild-type DNA using these two primers resulted in a 2.9-kb product that spanned the majority of the gene. This PCR reaction served to verify that the 5′ and 3′ primers were functioning properly. In addition, the resulting product was used as a hybridization probe later in the procedure.

The beginning and end of a T-DNA molecule are defined by specific DNA sequences called the left border and right border, respectively (1). Ideally, these L and R sequences are present at the junctions between the bacterial and plant DNA after the T-DNA has integrated into the plant chromosome. PCR primers specific for the left border and right border were therefore designed (Fig. 1*A*). In practice, it has been shown that *ca.* 50% of the T-DNA-transformed lines used in this study carry intact left border/plant junctions, whereas 25% carry intact right border/plant junctions (5). Taken together, *ca.* 75% of the *Arabidopsis* lines contain a junction between plant DNA and either the right or the left border.

To identify a T-DNA insertion within the CPK-9 gene, four different combinations of PCR primers can be used (5′ + L, 5′ + R, 3′ + L, and 3′ + R) (Fig. 1*A*). As diagrammed in Fig. 1*A*, none of these primer combinations will produce PCR products when wild-type DNA is used as a template. By contrast, the CPK-9 gene carrying a T-DNA insert as shown gives a 2.8-kb PCR product with primers 5′ + R and a 300-bp product with primers 3′ + L. The appearance of a PCR product in a reaction using one gene-specific primer and one

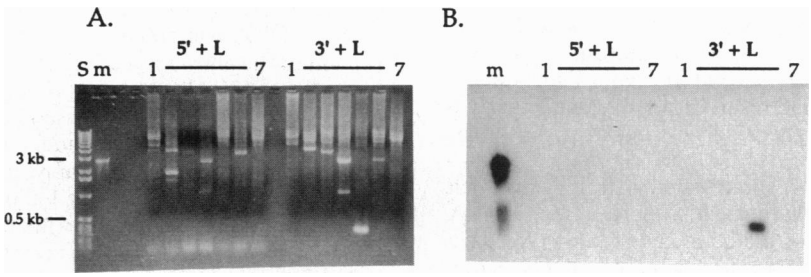


FIG. 2. PCR using DNA from the seven super pools. (*A*) PCR was performed using two pairs of primers (5′ + L) and (3′ + L), and the products were resolved by agarose gel electrophoresis. Lanes 1–7 correspond to the seven super pools. S, 1-kb DNA ladder size standard; m, 2.9-kb product produced by amplification of the wild-type CPK-9 gene using primers 5′ + 3′. (*B*) The agarose gel shown in *A* was Southern blotted and hybridized using the 2.9-kb CPK-9 product as a probe. The resulting autoradiogram shown in *B* shows a strong hybridization signal in the super pool 5 lane which was amplified using primers 3′ + L.

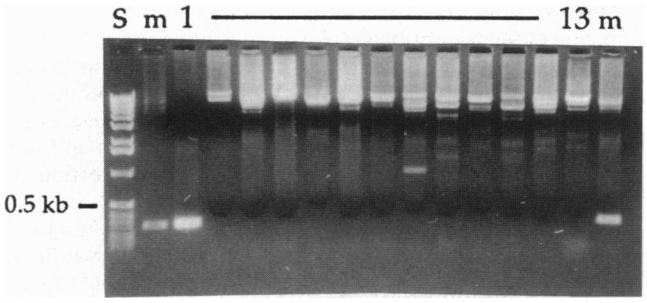


FIG. 3. PCR using DNA from the pools of 100. DNA from the 13 pools of 100 that make up super pool number 5 were subjected to PCR using the primers 3′ + L. Lanes 1–13 correspond to these 13 pools of 100. S, 1-kb DNA ladder size standard; m, 300-bp CPK-9/T-DNA product produced by amplification of super pool 5 using primers 3′ + L.

T-DNA-specific primer is therefore indicative of a T-DNA insertion in or near the gene of interest. As shown below, it is also necessary to perform a Southern blot and hybridization with the gene of interest to confirm the identity of a particular PCR product.

Organization of the T-DNA-Transformed *Arabidopsis* Lines. Our study uses Kenneth Feldmann's T-DNA-transformed *Arabidopsis* lines (2, 6). Of these independently transformed lines, 6400 are available through the ABRC as pools of 100 or in constituent pools of 20. A pool of 100 is a batch of seed that is derived from 100 independently transformed parental lines. An additional 2700 lines organized only as pools of 100 were also available to us from the DuPont collection. The overall organization of the seed pools used in our study is diagrammed in Fig. 1*B*. It has been shown that these *Arabidopsis* lines generated by seed transformation have a T-DNA inserted at an average of 1.5 loci per transformed line (2, 6). This collection of 9100 plants therefore represents *ca.* 13,650 different T-DNA inserts.

Seeds from each of the 91 pools of 100 were grown in liquid culture, and genomic DNA was isolated from the resulting plant material. Portions of these 91 DNA preparations were then combined to form seven super pools, with each super pool representing 13 or 14 pools of 100 (Fig. 1*B*). By combining the DNA preparations into seven super pools, we were able to greatly reduce the number of PCR reactions necessary to screen for a T-DNA insert in a particular gene. Control experiments using a previously identified T-DNA insertion had indicated that our PCR conditions were sensitive enough to easily detect one T-DNA-tagged line against a complex background of 1300 lines containing other T-DNA inserts (data not shown).

PCR Using the Seven Super Pools. The first step toward identifying a T-DNA insert within a particular gene is to subject the seven super pools to PCR reactions using 5′ and 3′ gene-specific primers in combination with the T-DNA border-specific primers (Fig. 1*A*). These PCR reactions are then run on an agarose gel, stained with ethidium bromide, and photographed. The result obtained using CPK-9-specific primers in conjunction with the T-DNA left border primer is shown in Fig. 2*A*. As one can see, there are many PCR products visible on the agarose gel. We have found that the large majority of these products are the result of amplification primed on both sides by the T-DNA primer alone (data not shown). Since T-DNA elements are known to be present in complex, rearranged forms in some of the transformed *Arabidopsis* lines, it is expected that there are templates in this population of 9100 lines that can be amplified by a single T-DNA-specific primer.

Given the presence of spurious PCR products in these reactions, it is necessary to perform Southern blotting of the agarose gel and subsequent hybridization using wild-type CPK-9 DNA as a probe (Fig. 2*B*). Only those PCR products with high homology to the CPK-9 gene will be detected on the Southern blot, thereby eliminating the background problem caused by the numerous other products. As seen in Fig. 2*B*, amplification of super pool 5 using the primer pair 3′ + L gave rise to a single 300-bp product with high homology to the CPK-9 probe. This result strongly suggested that super pool 5 contained genomic DNA that came from a plant carrying a T-DNA within the CPK-9 gene.

The data obtained so far show that a CPK-9-homologous PCR product can be generated from super pool 5 using primers 3′ + L. To confirm that these products were the result of an authentic T-DNA insertion in the CPK-9 gene, we found it helpful to perform a second round of PCR using a nested T-DNA primer. In particular, the products from the 3′ + L reaction were diluted and used as the templates in a new PCR reaction. This second reaction was run using the same 3′ primer and a new left border primer, which hybridized to T-DNA sequences interior to the original left border primer. This test will rule out the possibility that the PCR product of interest has been generated from a wild-type CPK-9 template due to fortuitous priming by the left border primer. The 300-bp product shown in Fig. 2*A* passed this test (data not shown).

PCR Using Pools of 100. We have shown above that super pool 5 tested positive for a T-DNA insert within the CPK-9 gene. The next step was to determine which particular pool of 100 contains this insert. The 13 pools of 100 corresponding to super pool 5 were therefore subjected to PCR using the primer pair 3′ + L. As seen in Fig. 3, a single pool of 100 was found to produce the expected 300-bp PCR product. This experiment demonstrated that the ABRC pool of 100 CS6494 contains seeds carrying a T-DNA inserted within the CPK-9 gene.

PCR Using Pools of 20. Genomic DNA was isolated from the five ABRC pools of 20 corresponding to the pool of 100 CS6455. PCR reactions were then performed using the primer pair 3′ + L. As seen in Fig. 4, a single pool of 20 was found to produce the expected 300-bp product. This experiment demonstrated that the ABRC pool of 20 CS6467 contains seeds carrying a T-DNA inserted within the CPK-9 gene.

Finding the Individual Plant. Our final goal was to isolate an individual plant that carries the T-DNA inserted within the CPK-9 gene. Toward this end, approximately 100 seeds from the ABRC pool of 20 CS6467 were planted in soil. Three weeks later, small sections of leaf tissue were removed from each plant, and genomic DNA was isolated from the tissue. For the first round of DNA preparations, leaf tissue was consolidated

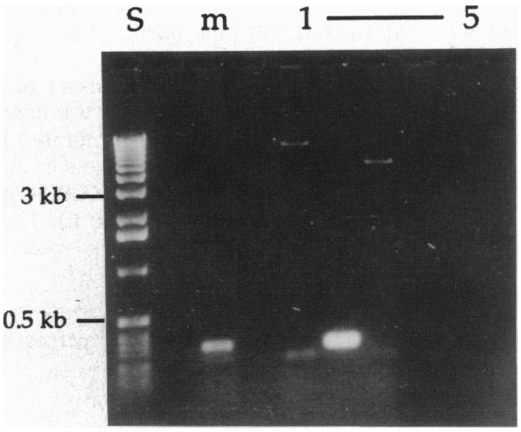


FIG. 4. PCR using DNA from the pools of 20. The 5 pools of 20 that make up pool of 100 number CS6494 were subjected to PCR using the primers 3′ + L. Lanes 1–5 correspond to these 5 pools of 20. S, 1-kb DNA ladder size standard; m, 300-bp CPK-9/T-DNA product produced by amplification of pool of 100 number CS6494 using primers 3′ + L.

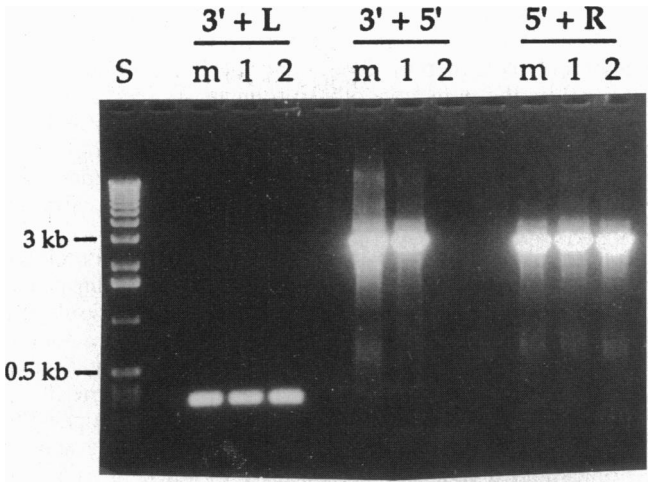


FIG. 5. PCR on DNA extracted from individual plants. DNA isolated from two individual plants found to carry the CPK-9 T-DNA insert was subjected to PCR using various primers. Lanes 1 and 2 correspond to these two individual plants. m, amplified using DNA from pool of 100 number CS6494, which contains the CPK-9 T-DNA insert; S, 1-kb DNA ladder size standard.

into 10 pools of 10 plants each. This strategy allowed us to more efficiently screen a large number of plants. PCR reactions were performed using the primer pair 3′ + L. The desired 300-bp PCR product was detected in 3 of these pools of 10 plants. Genomic DNA was then isolated from individual plants and subjected to PCR to find the particular plants carrying the desired T-DNA insert. From the 100 seeds planted in soil, three plants that carried a T-DNA insert in CPK-9 were found.

Additional PCR reactions were then performed on these individuals to determine if the plants were homozygous or heterozygous for the insert. A plant homozygous for the T-DNA insertion should not generate products when amplification is performed using the 3′ + 5′ primers. Since a single T-DNA element is ≈14 kb in length and most insertions contain multimers of the repeat unit, the PCR conditions used in this study did not allow for the formation of such large products (2, 6). By contrast, a plant heterozygous for the T-DNA insert will have one wild-type copy of the CPK-9 gene, which will be amplified when PCR is performed using the 3′ + 5′ primers. As seen in Fig. 5, both a homozygote (lane 2) and a heterozygote (lane 1) were found. We also determined that PCR using the primer pair 5′ + R gave rise to a 2.8-kb product (Fig. 5). A product of this size is expected if the T-DNA is inserted with both the left and right borders intact, as shown in Fig. 1*A*.

DNA Sequence Analysis of the T-DNA Insert in CPK-9. DNA sequence of the T-DNA insert in CPK-9 was determined by directly sequencing the PCR products generated in reactions using the primer pairs 5′ + R and 3′ + L. Junctions between the T-DNA border sequences and the CPK-9 gene are seen for both the right and left border (Fig. 6). This result demonstrates the authenticity of the T-DNA insertion and shows that the T-DNA is inserted within the 3′ untranslated portion of the transcribed region of the CPK-9 gene. It was also noted that 50 bp of unknown sequence were present between the T-DNA right border and the CPK-9 gene. In addition, 21 bp of the CPK-9 gene were apparently deleted as a result of the T-DNA insertion. The presence of small deletions and unexpected DNA sequences at the sites of T-DNA insertions has been documented by others (3, 6).

Additional Genes Surveyed. We applied the technique described above to search for T-DNA insertions within the members of gene families thought to be involved in signal transduction and ion transport. To date, we have found 17 T-DNA inserts within the 63 genes surveyed. The particular genes analyzed and the locations of the T-DNA inserts are shown in Fig. 7. The identity of the genes targeted by each of these 17 T-DNA inserts was confirmed by DNA sequence analysis (data not shown). Our approach has been to first screen through a set of genes using the left border T-DNA primer. If no knockouts are found using the left border primer, then the procedure is repeated using the right border T-DNA primer. This strategy was adopted in light of the analysis by Castle *et al.* (5) of 36 Feldmann T-DNA lines that showed 23 left border/plant DNA junctions and only 9 right border/plant junctions.

DISCUSSION

We have presented a technique that allows one to quickly screen for T-DNA insertions within any sequenced *Arabidopsis* gene. Other laboratories interested in this technique are encouraged to use the CPK-9 T-DNA insert described in this paper as a positive control to verify that the method is functioning properly in their hands. One should be able to easily identify the diagnostic 300-bp PCR product when reactions are done using DNA prepared from the ABRC pool of 100 CS6494.

As molecular biologists characterize more and more DNA sequences from *Arabidopsis thaliana*, it is becoming apparent that many processes are under the control of multigene families (21). With these multigene families comes the potential for redundant functions being provided by the related members of a particular gene type. This situation presents a challenge for the scientist trying to understand the functions of these genes *in vivo*. Simple genetic analysis may not uncover mutants in a particular gene if a redundant function is being provided by a related gene. One solution to this problem is to use reverse genetics and systematically "knockout" all of the members of a gene family one at a time. If no phenotype is discernible in the single gene knockout, one can use genetic crosses to create double mutants carrying knockouts of two different members of the gene family. One could continue to experiment with all possible combinations of double mutants, triple mutants, etc., until a phenotype is uncovered. The technique described in this paper should be useful for scientists interested in assembling such collections of gene knockouts.

Using this technique, we were able to locate T-DNA insertions in 25% of the genes surveyed (16 out of 63). Our

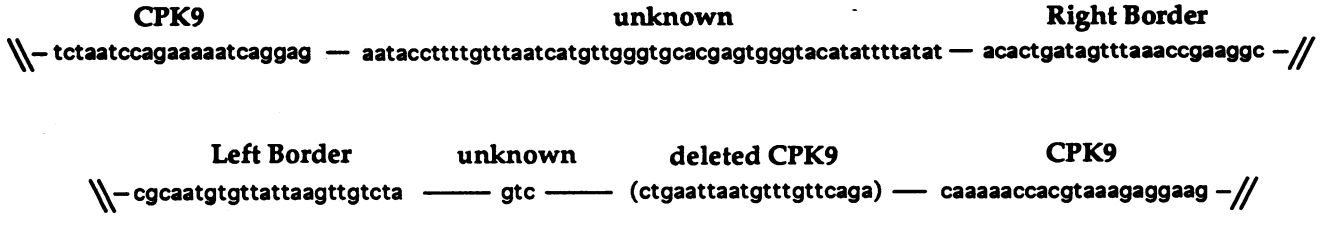
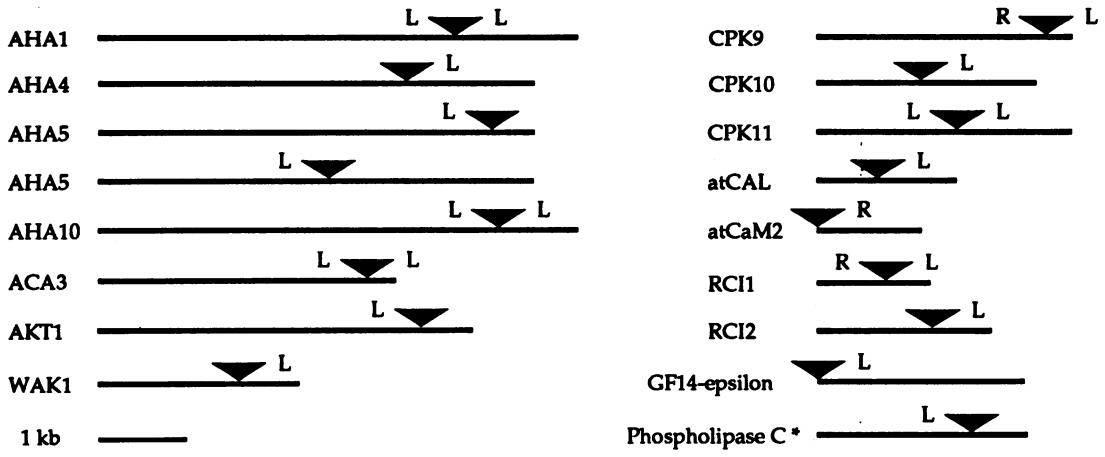


FIG. 6. DNA sequence of the T-DNA insertion site in CPK9. The PCR products shown in Fig. 5 were directly sequenced using the right and left border primers to read sequence out of the T-DNA into the flanking genomic DNA. Unknown, DNA sequence with homology to neither the CPK9 gene nor the T-DNA.



No inserts found: AHA2, AHA3, AHA6, AHA7, AHA8, AHA9, AHA11, AHA13, AHA14, ACA1, ACA2, AMA1, CPK1, CPK2, CPK4, CPK5, CPK7, CPK8, CPK12, atCAM, atCAM1, atCAM3, atCAM4, GRF2, GF14-upsilon, GF14-phi, GF14-kappa, ATMPK1, ATMPK2, ATMPK3, ATMPK4, ATMPK5, ATMPK6, ATMPK7, ATHATPLC1, ATHATPLC2, KAT1, RLK1, RLK4, RLK5, ARK1, TMK1, TMK4, ERS, ETR1, ETR2, AGL15. * Genbank ATU13203.

FIG. 7. Summary of our search for T-DNA inserts in 63 different genes involved in signal transduction and ion transport. Triangles show the approximate locations of the T-DNA inserts for each gene. L and R designate T-DNA left and right border sequences that serve as templates for the L and R primers. The absence of an L or R from one side of a triangle indicates that we have not characterized that side of the T-DNA/plant junction. The gene abbreviations are as follows: AHA, H^+ ATPases (7); ACA, Ca^{2+} ATPases; AMA, metal pump; CPK, calmodulin domain protein kinases (3); atCAL and atCAM, calmodulins (8); GRF, GF14, and RCI, 14-3-3 related proteins (9, 10); ATMPK, map kinases (11); ATHATPLC, phospholipase Cs (12); AKT and KAT, K^+ channels (13, 14); RLK, ARK, TMK, and WAK, receptor kinases (15–17); ERS and ETR, ethylene receptors (18, 19); AGL, MADS box protein (20). *, GenBank accession no. ATU13203.

calculations suggest that this figure is near the expected success rate. To begin with, the mean size of the 63 genes analyzed was 3.2 kb. In the terms of this study, we can therefore think of the 100,000-kb *Arabidopsis* genome as being made up of 31,250 different 3.2-kb targets. Now we may ask how many of these 31,250 targets should have T-DNA hits in the population of 9100 transformed lines. Since there are T-DNA inserts at an average of 1.5 loci per transformed line, we actually have a population of 13,650 independent T-DNA insertions. Of these 13,650 insertions, however, only 75%, or 10,238, have a left and/or right border/plant DNA junction that would be detectable with our PCR primers. One can then use the formula $P = 1 - (1 - f)^n$ to calculate the probability (P) that any given gene has an insert, where $f = 1/$(the number of targets) and $n =$ the number of T-DNA inserts (22). Given 10,238 potentially detectable T-DNA inserts in a genome made up of 31,250 targets, we find that any given gene should have a 28% chance of having a T-DNA insert. Our observed success rate of 25% suggests that our technique has the sensitivity to detect a large majority of the T-DNA inserts present in the population.

Other laboratories interested in searching for knockout mutants of their favorite genes must bear in mind that any particular gene only has a 25% chance of containing a T-DNA insert, given the currently available population of transformed lines. One therefore wants to be very confident that a negative result is due to the absence of that mutant in the population rather than a shortcoming of the method. The demonstration above that our observed success rate matches the expected success rate of 25% is direct experimental evidence of our method's reliability. It should be noted that we optimized the DNA preparation procedure and PCR conditions before we embarked on our survey of 63 different genes. These preliminary control experiments allowed us to develop the level of sensitivity and reproducibility needed to ensure that the failure to detect a knockout plant reflects its absence in the population, rather than a fault in the screening technique. By closely following the procedures that we describe in this paper, other researchers can benefit from this groundwork and embark on their own search for knockout mutants.

With the 9100 T-DNA transformed lines available for this study, we identified T-DNA inserts in 25% of the genes surveyed. Given this success rate, how many more T-DNA transformed lines would have to be analyzed to approach saturation of the genome with T-DNA hits? Calculations using the formula described above indicate that a total of 95,000 T-DNA transformed lines would be required to have a 95% chance of hitting every 3.2-kb gene (22). A more realistic goal of finding an insert in 50% of *Arabidopsis* genes would require a total of 22,000 lines. With a new batch of 5400 lines recently made available from the collection at DuPont along with the 9100 lines described in this study, the success rate of this method could improve to *ca.* 37%, or one in three randomly surveyed 3.2-kb genes.

The authors wish to thank Leslie Dickmann for DNA sequencing, Rebecca Hirsch for providing data on the K^+ channel knockout, Rob Meister for technical assistance, Estelle Hrabak and Jeff Harper for DNA sequences, and Richard Meagher for providing a preprint of his manuscript (3). This work was supported by a National Science Foundation (NSF) Postdoctoral Fellowship in Plant Biology awarded to P.J.K. and grants from the U.S Department of Energy/National Science Foundation/U.S. Department of Agriculture Collaborative Program in Plant Biology (BIR 92-20331), the U.S. Department of Energy (DE-F602-88ER13938), and the U.S. Department of Agriculture National Research Initiative (Competitive Grant 92-373048058).

1. Zambryski, P. (1988) *Annu. Rev. Genet.* **22,** 1–30.
2. Feldmann, K. A., Malmberg, R. L. & Dean, C. (1994) in *Arabidopsis*, eds. Meyerowitz, E. M. & Somerville, C. R. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 137–172.
3. McKinney, E. C., Ali, N., Traut, A., Feldmann, K. A., Belostotsky, D. A., McDowell, J. M. & Meagher, R. B. (1995) *Plant J.* **8,** 613–622.
4. Hrabak, E., Dickman, L., Saterlee, J. S. & Sussman, M. R. (1996) *Plant Mol. Biol.*, in press.
5. Castle, L. A., Errampalli, D., Atherton, T. L., Franzmann, L. H., Yoon, E. S. & Meinke, D. W. (1993) *Mol. Gen. Genet.* **241,** 504–514.
6. Feldmann, K. A. (1991) *Plant J.* **1,** 71–82.

7. Sussman, M. R. (1994) *Annu. Rev. Plant Phys. Plant Mol. Biol.* **45,** 211–234.
8. Ling, V., Perera, I. & Zielinski, R. E. (1991) *Plant Physiol.* **96,** 1196–1202.
9. Ferl, R. J., Lu, G. & Bowen, B. W. (1994) *Genetica* **92,** 129–138.
10. Jarillo, J. A., Capel, J., Leyva, A., Martinez-Zapater, J. M. & Salinas, J. (1994) *Plant Mol. Biol.* **25,** 693–704.
11. Mizoguchi, T., Hayashida, N., Yamaguchi-Shinozaki, K., Kamada, H. & Shinozaki, K. (1993) *FEBS Lett.* **336,** 440–444.
12. Hirayama, T., Ohto, C., Mizoguchi, T. & Shinozaki, K. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 3903–3907.
13. Sentenac, H., Bonneaud, N., Minet, M., Lacroute, F., Salmon, J. M., Gaymard, F. & Grignon, C. (1992) *Science* **256,** 663–665.
14. Anderson, J. A., Huprikar, S. S., Kochian, L. V., Lucas, W. J. & Gaber, R. F. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 3736–3740.
15. Walker, J. C. (1993) *Plant J.* **3,** 451–456.
16. Tobias, C. M., Howlett, B. & Nasrallah, J. B. (1992) *Plant Physiol.* **99,** 284–290.
17. Chang, C., Schaller, G. E., Patterson, S. E., Kwok, S. F., Meyerowitz, E. M. & Bleecker, A. B. (1992) *Plant Cell* **4,** 1263–1271.
18. Chang, C., Kwok, S. F., Bleecker, A. B. & Meyerowitz, E. M. (1993) *Science* **262,** 539–544.
19. Hua, J., Chang, C., Sun, Q. & Meyerowitz, E. M. (1995) *Science* **269,** 1712–1714.
20. Rounsley, S. D., Ditta, G. S. & Yanofsky, M. F. (1995) *Plant Cell* **7,** 1259–1269.
21. Pickett, F. B. & Meeks-Wagner, D. R. (1995) *Plant Cell* **7,** 1347–1356.
22. Clarke, L. & Carbon, J. (1976) *Cell* **9,** 91–99.