The Arabidopsis Knockout Facility at the University of Wisconsin–Madison[1]

Michael R. Sussman*, Richard M. Amasino, Jeffery C. Young, Patrick J. Krysan, and Sandra Austin-Phillips

Biotechnology Center (M.R.S., P.J.K., S.A.-P.) and Department of Biochemistry (R.M.A.), University of Wisconsin, Madison, Wisconsin 53706; and Biology Department, Western Washington University, Bellingham, Washington 98225 (J.C.Y.)

As of this writing the Arabidopsis genome is 97% sequenced with only small portions of the highly repetitive regions within centromeres and telomeres remaining. The identification of approximately 25,000 plant genes will give plant biologists an opportunity to identify and understand the function of the proteins they encode. One exciting tool that will aid in this endeavor is the use of insertional mutagenesis to create "gene knockouts." The availability of a mutant line in which the action of a known, specific gene has been disrupted gives the plant biologist a powerful tool in understanding the action of that gene. The basis of this approach is to create a large population of plants containing randomly inserted pieces of foreign DNA. If the sequence of a gene is known, it is possible to devise a PCR-based strategy to identify a plant where that specific gene has been disrupted by the insertion of foreign DNA. To fully utilize this technology it is necessary to saturate the genome with insertion mutations and to develop efficient PCR-based screening methods to comb through knockout plant populations and identify specific mutant plants. The smaller the gene, the more difficult a target it represents, and thus hundreds of thousands of lines are needed to provide a high probability that a particular gene is present as a knockout in the population.

As a beginning, a method has been developed for rapidly searching a large collection of T-DNA transformed Arabidopsis lines for the presence of T-DNA inserts within specific genes (Krysan et al., 1996, 1999). To use this technology, a collection of 60,480 Arabidopsis (accession Wassilewskija [WS]) lines were generated that were transformed with the T-DNA vector pD991. Preliminary data from screening this collection indicated that this collection could indeed be efficiently screened for mutant lines. To share this resource with all members of the Arabidopsis research community, a "Knockout Facility" was established at the University of Wisconsin (Madison) in 1999 as part of the Arabidopsis Functional Genomics Consortium. A detailed description of this initial population and the operation of the facility have been given in a recent publication (Krysan et al., 1999) and is described more fully at the website http://www.biotech.wisc.edu/Arabidopsis.

In this brief note we will give an overview of the first year's operation and describe our future plans. We strongly advice that users read the web site fully before using the facility and contact us with any questions that they may have.

CURRENT FACILITY OPERATION

The facility is housed in the Plant Biotechnology Laboratory at the University of Wisconsin Biotechnology Center located at 425 Henry Mall, University of Wisconsin, Madison, WI 53706. The fee-for-service operation relies heavily on the administrative and information services of the Biotechnology Center. Users interact with the facility through the website. The site has a full description of the PCR screening including primer design and subsequent analysis of the PCR products. Primer design is a critical element of the screening process and users are asked to test their primers before sending them to the facility. Each user has their own personal web page on the site allowing them to track their screens. Our current resources are organized as follows:

Seed: 6,720 pools of 9
DNA Preps: 270 pools of 225
DNA Super Pools: 30 super pools (9 pools of 225 per super pool; 2,025 lines per super pool)

Researchers basically send us PCR primers for a gene for which they want a "knock-out." We perform PCR reactions using our DNA pools and the researcher's PCR primers. We then send the PCR reactions to the users to analyze. Our PCR strategy is composed of two rounds of PCR. The first round of PCR searches the entire population for T-DNA inserts in the gene. This set of reactions uses the researcher's gene primers and one T-DNA border primer and the 30 DNA super pools. If a hit is found in the first round of PCR then a second round is performed which narrows down the hit to one particular pool of 225. Last of all, the researchers request seed from the 25 pools of 9, which correspond to that pool of 225. One of these tubes of seed will contain the knockout. The researchers will then

[1] This work was supported by the National Science Foundation (grant no. DBI 9872638).

* Corresponding author; e-mail msussman@facstaff.wisc.edu; fax 608–262–6748

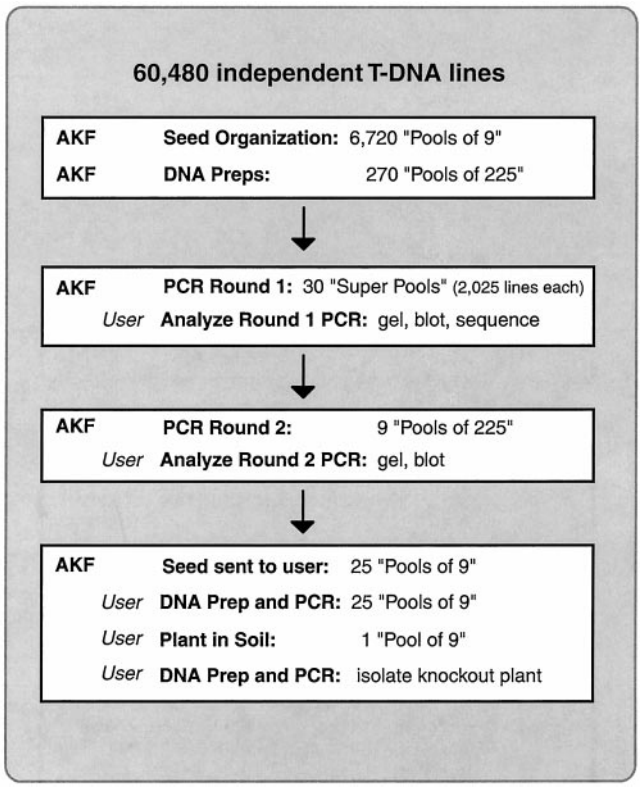

Figure 1. Organization and screening strategy for the first knockout population.

search within this population for the plant(s) in which their gene is disrupted. The organization and screening strategy is summarized in Figure 1.

RESULTS TO DATE

The laboratories of over 360 Principal Investigators are currently using the Knockout Facility; 60% of these are from North America, 25% are from Europe, and the remainder are from other parts of the world. As of August 2000, over 1,300 first-round and 900 second-round screens have been performed. About 60% of users have found at least one hit in their first round screens. Some users choose to follow more than one hit. It is still too soon to accurately assess the number of screens that ultimately result in mutant plants with a useful phenotype.

FUTURE PLANS

Additional Populations

Although the current population has successfully identified knockouts for many users, there is obviously a need to incorporate additional populations into the screening process. We have generated a further 72,960 BASTA (glufosinate)-resistant lines transformed with an activation-Tag vector, pSK1015 (Weigel et al., 2000). Full details of the vector used and the screening process are given on the web site. This population will be available at the beginning of 2001. Seed for these lines is organized as pools of 10 in a 96-well format and the DNA is organized in a three-dimensional grid. This allows us to localize a knockout to a single pool of 10 plants after only two PCR rounds. This is obviously much more efficient for the user who currently has to find their knockout from 25 pools of nine plants. The primer design for this population is identical. Additional populations developed at the University of Wisconsin, Yale University, and other locations will be incorporated into the facility as they become available. It should be noted that for this large-scale effort aimed at saturating the genome, useful populations require a minimum of many tens of thousands of individual lines organized in pools of less than 10 lines each.

User Feedback

Screens are currently confidential and we do not maintain a database of genes or sequences for which a knockout has been obtained. This obviously leads to duplication of effort, but not all researchers are comfortable with their research efforts being public before the work is published. We intend to establish a searchable database on our website where users can voluntarily list their primer sequences and the outcome of their screens. We will also develop a survey for user feedback that will give basic information on the actual insertion events.

Long-Term Goals

PCR screening has proven to be an effective method of identifying plant knockouts. An even more efficient alternative would be to catalog all of the locations of the T-DNA inserts within the population using a sequence tag. This database of T-DNA flanking sequences could be searched for the presence of flanking sequences homologous to any gene of interest, thus identifying a knockout plant.

Pilot studies are currently in progress to develop a high throughput procedure, e.g. using thermal asymmetric interlaced PCR to accomplish the sequencing. Efforts are now under way to analyze thousands of flanking sequences from individual lines. It will be necessary to have hundreds of thousands of such sequences to be a fully comprehensive database. Since this may take several years to create, we intend to keep the PCR-based screening facility as a parallel effort until the database of sequences is sufficiently large to ensure that researchers are certain to obtain knockout plants in their gene of interest. Researchers have also found it extremely useful to obtain more than one allele, i.e. knockouts created by insertions in various locations within the gene. Such multiple alleles are useful to be sure that resultant phenotypes are in fact due to the insertion rather than mutations located nearby, but not within the gene of interest.

Future populations will contain transposons located within the T-DNA to allow a line to be developed in which neighboring genes are disrupted. The current method of creating "double knockouts" involves first isolating individual knockout lines and then crossing them into the same plant. For genes that are located within a few thousand bases of each other this method would require the screening of an unreasonable number of recombinants. The alternative method of using the first insertion line as a "launching pad" for transposon elements to land nearby represents a more reasonable and viable approach.

ACKNOWLEDGMENTS

The authors thank Sean Monson, Laura Katers, Sarah Benn, Kiersten Iovinella, and Erin Olsen for technical assistance in the Knockout Facility.

Received September 1, 2000; accepted September 22, 2000.

LITERATURE CITED

Krysan PJ, Young JC, Sussman MR (1999) T-DNA as an insertional mutagen in *Arabidopsis*. Plant Cell **11:** 2283–2290

Krysan PJ, Young JC, Tax F, Sussman MR (1996) Identification of transferred DNA insertions within *Arabidopsis* genes involved in signal transduction and ion transport. Proc Natl Acad Sci USA **93:** 8145–8150

Weigel D, Ahn JH, Blazquez MA, Borevitz JO, Christensen SK, Fankhauser C, Ferrandiz C, Kardailsky I, Malancharuvil EJ, Neff MM, Nguyen JT, Sato S, Wang ZY, Xia Y, Dixon RA, Harrison MJ, Lamb CJ, Yanofsky MF, Chory J (2000). Activation tagging in Arabidopsis. Plant Physiol **122:** 1003–1013