

Reverse Genetics: Identification of Arabidopsis Lines with T-DNA Insertions in a Gene of Interest

Brenda G. Leicht¹ and Chi-Lien Cheng²

¹Department of Biology
168 Biology Building
The University of Iowa
Iowa City, IA 52242
brenda-leicht@uiowa.edu

²Department of Biology
210 Biology Building
The University of Iowa
Iowa City, IA 52242
chi-lien-cheng@uiowa.edu

Biography

Brenda G. Leicht is a Laboratory Coordinator and Adjunct Assistant Professor in the Department of Biology at The University of Iowa. She coordinates the labs and lectures for *Human Biology*, a large non-majors course, as well as coordinating two advanced investigative laboratory courses for majors, *Cell Biology Laboratory* and *Genetics and Biotechnology Laboratory*. She received her Ph.D. in Genetics from Indiana University in 1987. She has also taught in the Department of Biology at Pennsylvania State University.

Chi-Lien Cheng is Associate Professor of Biology and Graduate Director for the Department of Biology at The University of Iowa. She teaches Molecular Genetics and Genetics and Biotechnology Laboratory. She received her Ph.D in Biology from The University of Connecticut in 1982. Her current research focuses on alternation of generations in land plants and in vegetative phase change in maize.

Introduction

The completion of genome sequencing projects in a variety of model organisms has made it possible to use two complementary approaches to study gene function. The classical “forward genetics” approach is to generate random mutations that give a phenotype of interest and to use the mutant phenotype to map and subsequently sequence the gene. The newer “reverse genetics” approach starts with the gene sequence rather than the mutant phenotype. In the reverse genetics approach, a large collection of random insertions of mobile DNA elements is generated, ideally to yield at least one insertion in every gene. A researcher simply needs to identify the line(s) with the insertion in their gene of interest and subsequently analyzes the line for the phenotypic effects of the altered (i.e., disrupted) gene.

The small dicotyledonous plant *Arabidopsis thaliana* is particularly amenable to reverse genetic analysis. Its 125 Mb genome has been completely sequenced, and approximately 25,000 unique genes have been identified. Many of these genes were not previously analyzed by traditional forward genetics. As a first step in determining the function of all *Arabidopsis* genes, a consortium of *Arabidopsis* researchers has generated a large collection of T-DNA insertion mutants (Alonso et al., 2003). (T-DNA is a mobile piece of DNA carried by the Ti plasmid of the soil bacterium *Agrobacterium tumefaciens* and it can insert at random in the genome of infected plants.) The Salk Institute Genome Analysis Laboratory (SIGnAL) has been using high-throughput sequencing methods to identify the sites of the T-DNA insertions within the *Arabidopsis* genome. Following identification of the mutant locus in a line, the lines are made publicly available through the *Arabidopsis* Biological Resource Center (ABRC) at Ohio State University.

We have made use of the excellent resources of the *Arabidopsis* research community to demonstrate the use of reverse genetics to students taking a course in Genetics and Biotechnology. Students begin with a gene of interest (e.g., our students have previously searched for a gene encoding a co-chaperone protein and another that encodes an exportin protein) and use the TAIR and SIGnAL websites to find T-DNA insertions in that gene. They next use a primer design tool, also at the SIGnAL website, to retrieve the sequences of PCR primers that will give PCR products that distinguish the three possible genotypes (homozygotes for the T-DNA insertion, heterozygotes for the T-DNA insertion, and homozygotes for the wild type allele). The wet lab portion of the exercise involves growing the T-DNA insertion mutation lines (which are a genotypically mixed population of T₃ seeds), isolating genomic DNA from individual T₃ plants, and using PCR and agarose gel electrophoresis to determine the genotype (specifically those that are hetero- or homozygous for the T-DNA insertion) of the individual plants.

The basic design of this laboratory exercise could be modified and applied to the functional analysis of any gene of a model genetic organism for which the genome is sequenced and for which insertional mutants have been systematically identified. Several important instructional objectives are met by the lab exercise: development of website navigation skills and use of websites for research and experimental design purposes, understanding of the strategy of reverse genetics and ways to identify the desired mutant organisms, review of chromosome segregation (because the T-

DNA insertion lines are frequently heterozygous), and hands-on application of PCR to genetic analysis.

Due to time constraints, this mini-workshop focused on the use of web-based resources for the design of reverse genetics experiments in *Arabidopsis thaliana*. Information for performing the wet lab analysis is provided in the Instructor Notes.

Student Outline

Utilization of web-based resources for genetic analysis of a gene of interest

1. **Open the Internet browser on your computer.**
2. **Go to the TAIR website ([http:// www.arabidopsis.org/index.jsp](http://www.arabidopsis.org/index.jsp)).** The Arabidopsis Information Resource, or TAIR, maintains a comprehensive database of genetic and molecular biology data for *Arabidopsis thaliana*, including the complete genome sequence along with gene structure, gene product and expression information, DNA and seed stocks, genome maps, genetic and physical markers, publications, and news and information about the Arabidopsis research community. Functional data is updated every two weeks. Gene structures are updated once or twice per year.
3. **Use the Search function at the top of the TAIR homepage to find a gene of interest.** You can enter the name of the gene product or the systematic name based on its chromosomal location (e.g., *At3g05040* where *At* refers to *Arabidopsis thaliana*, *3* to the third chromosome, *g05040* to the genetic locus).
 - a. Two possible choices:
 - At2g15790* (cyclophilin 40, a chaperone protein)
 - At3g05040* (member of the importin/exportin family involved in miRNA transport)
 - b. Choose a protein/gene of interest to you.
4. Study the output and **select one or more of the genes to work on.** Record the systematic name(s) here for future reference.

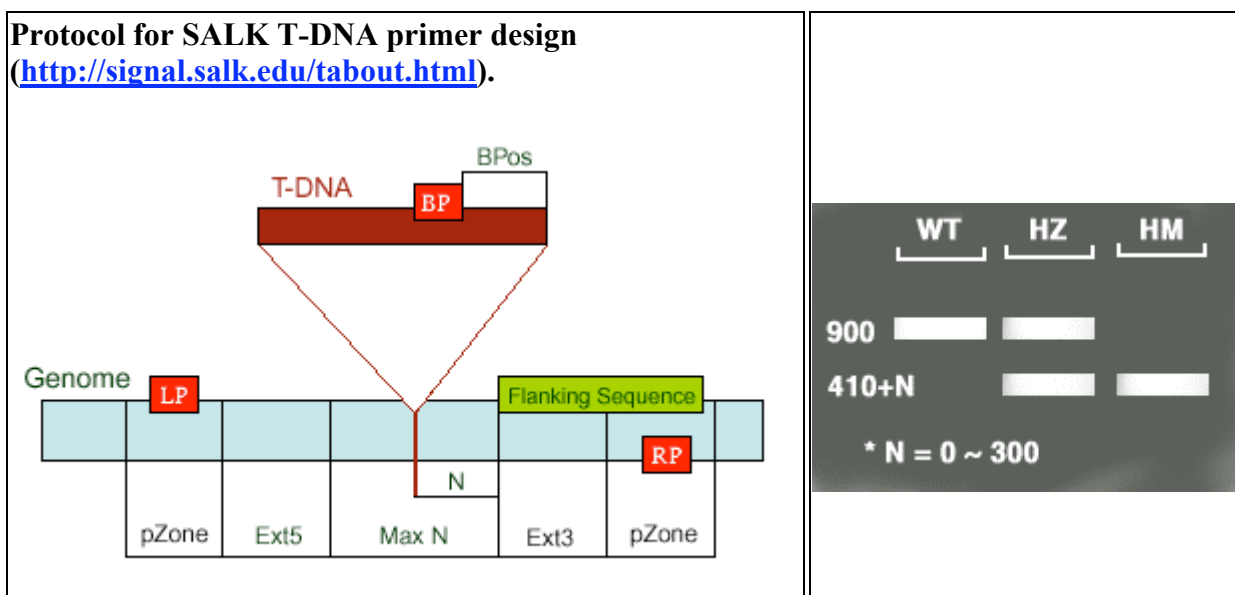
5. **Go to the SIGnAL website (<http://signal.salk.edu/about.html>).** Find **SALK T-DNA** in the main menu, then select **T-DNA Express** from the submenu. T-DNA Express will link you to the sequence-indexed library of T-DNA insertional mutations in the Arabidopsis genome that has been compiled by The Salk Institute Genome Analysis Laboratory (**SIGnAL**). SIGnAL has identified the specific sites of insertions in the genomes of the Alonso/Crosby/Ecker collection of T-DNA transformed plants. Insertion site sequences are aligned with the Arabidopsis genome sequence and gene annotations are added.
6. **Once you are in T-DNA Express,** scroll down to the **Search box** and where it says “**Query**”, **type in the gene name that you selected at step 4. Click on Search.** (If you selected two different genes, you will need to repeat the process separately for the second gene). View the

insertion map to find the specific insertional mutation(s) for your gene. Choose two or three insertion lines with the SALK prefix and record their numbers.

Gene 1 SALK Insertion Lines: _____, _____, _____

Gene 2 SALK Insertion Lines: _____, _____, _____

7. **Next open the iSect Tool page from the SIGnAL main menu** (to go back to the main menu click on the SIGnAL logo at the top of the page). **Click on 1. SIGnAL T-DNA Verification Primer Design**. At the top of the page is an explanation of how to use the appropriate primers to distinguish the three genotypes in a plant population (WT = homozygotes for the wild type allele, HM = homozygotes for the T-DNA insertion, and HZ = heterozygotes with both alleles). This page (with some minor edits) is copied below.



N - Difference of the actual insertion site and the flanking sequence position, usually 0 - 300 bases

MaxN - Maximum difference of the actual insertion site and the sequence, default 300 bps

pZone - Regions used to pick up primers, default 100 bps

Ext5, Ext3 - Regions between the MaxN to pZone, reserved not for picking up primers

LP, RP - Left, Right genomic primer

BP - T-DNA border primer LB - the left T-DNA border primer

BPos - The distance from BP to the insertion site

LB - Left border primer of the T-DNA insertion:

>[LBb1](#) of pBIN-pROK2 for SALK lines GCGTGGACCGCTTGCTGCAACT

>[Lba1](#) of pBIN-pROK2 for SALK lines TGGTTCACGTAGTGGGCCATCG

>LB_6313R for SALK lines TCAAACAGGATTTTCGCCTGCT

Using the three primers (LBb1+LP+RP) for SALK lines, WT (Wild Type - no insertion) should give a product of about 900-1100 bps (from LP to RP), HM (Homozygous lines - insertions in both chromosomes) will give a band of 410+N bps (from RP to insertion site 300+N bases, plus 110 bases from LBb1 to the left border of the vector), and HZ (Heterozygous lines - one of the pair chromosomes with insertion) will give both bands. The product size should be 200 base larger if using Lba1 instead

of LBb1. Use of this protocol requires similar TM values for all the LB, LP and RP primers.

You also can set up two paired reactions, LP+RP and LB+RP. You should get a product in the LP+RP reaction for WT or HZ lines or get a blank lane for HM lines, while you should get a band in the LB+RP for HM or HZ lines.

8. **Scroll down the iSect Tool page to 2. SALK T-DNA verification primer design.** In the box that says **Primer L**, type the names of the several T-DNA lines you identified in Step 6. (Check the format box if you would like to see the entire gene sequence as well as the sequences of the two primers.) **Click on Submit.** You will be given the sequence of two PCR primers, LP and RP (complementary to sequences that flank your gene of interest), as well as the expected sizes of the PCR products for the wild type and insertion alleles. An example of results for the gene *At3g05040*, insertion line SALK_079289 is shown below.

```
SALK_079289.31.25.x  PRODUCT_SIZE 1064

PAIR_ANY_COMPL 0.00  PAIR_3'_COMPL 0.00  DIFF_TM 0.06

LP ATTTTGTGCAGGGTATTGCTG  Len 21  TM 60.01  GC 42.86
SELF_ANY_COMPL 0.00  3'_COMPL 0.00

RP ATCGACTGCAAATTTTCCATG  Len 21  TM 59.95  GC 38.10
SELF_ANY_COMPL 0.00  3'_COMPL 0.00

Insertion chr3 1403047  BP+RP_PRODUCT_SIZE 448-748
```

Primers for your gene(s) of interest:

Gene 1: _____
 LP: _____
 RP: _____
 PCR Product for Mutant (Insertion) Allele: _____
 PCR Product for Wild Type Allele: _____

Gene 2: _____
 LP: _____
 RP: _____
 PCR Product for Mutant (Insertion) Allele: _____
 PCR Product for Wild Type Allele: _____

Instructor Notes

Protocols for Wet Lab Analysis

Ordering Seeds:

With the names of the insertion lines and the primer sequences in hand you are ready to move on the wet lab analysis of insertion lines of interest. Seeds from T-DNA insertion lines can be ordered from Arabidopsis Biological Resource Center (ABRC) at Ohio State University The ARBC website is: <http://www.biosci.ohio-state.edu/~plantbio/Facilities/abrc/index.html>.

Growing Arabidopsis Plants from Seeds:

If you have never worked with Arabidopsis before, the following are useful websites. The TAIR website also has many useful links.

<http://arabidopsis.info/protocols/grow.html>

<http://www.biosci.ohio-state.edu/pcmb/Facilities/abrc/handling.htm>

Preparing Genomic DNA from Arabidopsis:

The following protocol and recipes are from our undergraduate course lab manual. They are easy and quick.

1. Use the lid of a 1.5 ml microcentrifuge tube to clip off a young leaf from an Arabidopsis seedling. Use a separate tube for each plant that you want to extract DNA from.
2. Use a small plastic pestle (Research Products International Corp.) to grind the leaf; grind for ~ 15 seconds without buffer. Use a separate pestle for each plant.
3. Add 400 μ l of DNA Extraction Buffer (200 mM Tris-HCl, pH 7.5-8; 250 mM NaCl; 25 mM EDTA, 0.5% SDS) to each tube and vortex for ~ 5 seconds. (At this stage, samples can be kept at room temperature for up to 1 hour until all samples have been collected).
4. Centrifuge that tubes at full speed (>10,000 rpm) for 1 minute to pellet cellular debris.
5. Transfer 300 μ l of the supernatant from each tube into fresh, labeled 1.5 ml tubes. Avoid taking any of the debris from the first tube.
6. Add 300 μ l isopropanol to each tube, mix by inverting the tubes several times, and leave at room temperature for ~ 2 minutes.
7. Centrifuge the tubes at full speed for 5 minutes to pellet the DNA.
8. Remove all the supernatant and dry the pellets by incubating the open tubes at 37°C for several minutes. Do not overdry the pellets or it will be difficult to dissolve the DNA.
9. Add 100 μ l of sterile TE (10 mM Tris-HCl, pH 7.5-8; 1 mM EDTA) or ddH₂O to each pellet and dissolve by gentle flicking of the tube (do not vortex).
10. The DNA is of suitable quality for digestion with restriction enzymes and for PCR. Freeze unused portions until needed.

Thermocycler Program:

The annealing temperature is dependent on the GC content (T_m) of the primers and should be tested prior to implementing the lab with students. However, the following program has worked in our hands for several different primer mixes.

1X: 5 min @ 94 C

40X: 30 sec @ 94 C

30 sec @ 60 C

1 min @ 72 C

1X: 5 min @ 72 C

Gel Electrophoresis:

The percentage of agarose to use depends on the expected sizes of the PCR products. Most products range in size from 500 bp – 1200 bp. A 1.5% agarose gel (made up in either TAE or TBE buffer) works well for separating DNA molecules in this size range.

Recipe for 10X TAE (1 L)

48.4 g Tris Base

11.4 ml Glacial Acetic Acid

20.0 ml 0.5 M EDTA, pH 8

Bring to 1 L with dH₂O

Recipe for 5X TBE (1 L)

60.5 g Tris Base

30.0 g Boric Acid

3.7 g EDTA

Bring to 1 L with dH₂O

Literature Cited

Alonso, José M, Anna N. Stepanova, Thomas J. Leisse, Christopher J. Kim, Huaming Chen, Paul Shinn, Denise K. Stevenson, Justin Zimmerman, Pascual Barajas, Rosa Cheuk, Carmelita Gadrinab, Collen Heller, Albert Jeske, Eric Koesema, Cristina C. Meyers, Holly Parker, Lance Prednis, Yasser Ansari, Nathan Choy, Hashim Deen, Michael Geralt, Nisha Hazari, Emily Hom, Meagan Karnes, Celene Mulholland, Ral Ndubaku, Ian Schmidt, Plinio Guzman, Laura Aguilar-Henonin, Markus Schmid, Detlef Weigel, David E. Carter, Trudy Marchand, Eddy Risseeuw, Debra Brogden, Albana Zeko, William L. Crosby, Charles C. Berry, and Joseph R. Ecker. 2003. *Genome-Wide Insertional Mutagenesis of Arabidopsis thaliana*. *Science*, 301: 653-657.