

## Creation of a Gene Database from Potato as a Teaching Tool in an Undergraduate Course in Molecular Biology

Michael A. Campbell  
School of Science  
Penn State Erie-The Behrend College  
Station Road  
Erie, PA 16563

Ph: 814-898-6474  
Email: mac17@psu.edu

Teaching Molecular Biology in an undergraduate curriculum has changed dramatically in the last few years. Most of the techniques used in molecular biology used to be reserved for large laboratories with big budgets, highly trained technicians, post-docs, and graduate students. Recent changes in technology have enabled many institutions to apply sophisticated experiments to the teaching of molecular biology in the undergraduate curriculum.

The Behrend College is a four-year undergraduate institution affiliated with Penn State University. The college offers a bachelors degree in the biological sciences and does not currently have a graduate program in this discipline. The program in biology is composed of 7 full time faculty, 1 part time instructor, and 1 full time technician. There are currently about 120 majors at the junior and senior level. The program offers to juniors and seniors a non-required course in molecular biology. This class meets for two 50-minute lectures periods and one three-hour laboratory per week. The class is heavily weighted on the laboratory and analytical techniques common to the field of molecular biology.

In order to develop student laboratory skills we feel a project-based approach is ideal. In order to develop this theme we have established a gene database that is central to the laboratory portion of the class. This database is a collection of individual gene sequences plus any expression studies, similarity searches, or references pertinent to the individual genes described. The construction of this database is the responsibility of the students enrolled in the molecular biology course. During the course of the semester the students are required to isolate a gene, sequence the gene, perform a sequence similarity of the gene sequence to the world wide databases at NIH and EMBL, generate a literature review, and construct a proposal outlining future research. This information is then deposited to the database for public availability.

The starting point for this exercise is a cDNA library constructed from mid-log phase suspension cultures of potato. This library is an amplified Lambda-Zap vector (Stratagene Cloning Systems) that originally contained about  $1 \times 10^6$  independent mRNA molecules. The library was amplified to  $1 \times 10^{13}$  clones which may have reduced the presence of some rare genes but out of 48 student projects no clones have been isolated more than once. The laboratory section of the course is designed for a 15-week semester.

## **Week 1: Introduction to the laboratory.**

Students were introduced to the standard laboratory rules and procedures associated with a molecular biology laboratory. This included how solutions were handled to maintain aseptic conditions and the proper procedures for the disposal of laboratory materials.

The students were placed into groups and each group was required to make stock solutions such as 1 M Tris (pH 8.0), 0.5 M EDTA (pH 8.0), Top agar (1 M Tris pH 8.0, 0.5 M EDTA, 5 M NaCl), and TE (10mM Tris (pH 8.0), 1 mM EDTA) (see Appendix A). The students were given a scale and the bottles of the chemicals from the stock room. They were required to make all necessary calculations and pH the solutions on their own. This exercise was intended to reinforce the concepts of molarity and pH, which are critical for success in a molecular biology laboratory.

The cDNA library was presented to the students the first week of class as plaques on a plate. Each student was required to use a sterile micropipet tip to pick three individual plaques and place the pick into 200 ul of SM buffer (see Appendix A) followed by the addition of 0.3% chloroform. Students were then required to make serial dilutions of the plaque picks in order to determine the number of plaque forming units present. This exercise was intended to develop student skills with the pipet, which were critical throughout the rest of the course. It was also intended to introduce them to bacteriophage as a tool in molecular biology. The plaque picks were allowed to equilibrate at 4 C until the next week's laboratory.

The isolation of three different plaque picks was chosen in order to prevent major set backs due to failure to isolate DNA or to generate sequence information from a given phage pick. This did create the potential for students to isolate and characterize more than one clone but time and monetary constraints prevented this and only one clone was processed for DNA sequencing.

## **Week 2: Determination of phage titer**

Students were required to determine the number of plaque forming units by infecting *E. coli* cells with phage from serial dilutions of their plaque picks. A 5ul aliquot of the phage dilutions was mixed with 300 ul of host cells (see Appendix B). The cells were incubated at 37 C for 15 min. Three ml of top agar was added and then the solution was poured onto pre-warmed NZY plates (Appendix B). The cultures were grown overnight at 37 C and the students were expected to remove the plates from the incubator the next day and calculate the number of plaque forming units present in the plaque picks.

### Week 3: Development of Phagemid

In a 50-ml sterile conical tube students combined the following:

200 $\mu$ L of the XL-1 Blue-MRF' at an OD600 =1.0 (Appendix B)
250 $\mu$ L of phage stock (containing $>1 \times 10^5$ phage particles)
1 $\mu$ L of ExAssist helper phage ( $>1 \times 10^{10}$ pfu/ml)

The mixture was incubated at 37° C for 15 minutes.

Three ml of LB broth (25 ml of LB broth for mass excision) was added and the tube was incubated for 2-2.5 hours at 37°C with shaking. There was some flexibility in the incubation time here since single-clone excision reactions can be safely performed overnight because clonal representation is not relevant. So if laboratory time is limited the students could run the reactions overnight and plate phagemids the following day.

The tubes were then heat to 70°C for 15 minutes and then centrifuged at 4000 x g for 15 minutes. The supernatant containing excised phagemid packaged as filamentous phage particles was transferred to a sterile tube and stored at 4°C.

The excised phagemids were plated by adding 200 $\mu$ L of freshly grown SOLR cells (Appendix B) to two 1.5mL tubes, adding 100 $\mu$ L of the phagemid supernatant to one tube, and 1 $\mu$ L of the phagemid supernatant to the other tube. The tubes were then incubated at 37°C for 15 minutes and then 200 $\mu$ L from each tube was spread onto LB-ampicillin agar plates and incubated overnight at 37°C. Colonies composed of bacteria containing identical plasmids, and therefore the identical gene from potato, were visible the next morning.

### Week 4: Isolation of DNA from a colony

The DNA isolation procedure was a modification of the alkaline lysis procedure describe by Sambrook et al. (1989). A colony was used to inoculate 10 ml of Terrific Broth (Appendix A) supplemented with 50 ug/ml of ampicillin. The culture was incubated overnight at 37 C. The next day a 1.5 ml aliquot of the culture was placed in a microfuge tube and centrifuged at 12,000 x g for 1 min. The medium was removed and the cellular pellet was resuspended in 200 ul of buffer (50mM glucose, 25mM Tris (pH 8.0), 10 mM EDTA). This was followed by the addition of 300 ul of freshly prepared 0.2 M NaOH, 1 % SDS. The tube was mixed by gentle inversion and incubated on ice for 5 min. A 300 ul aliquot of 3.0 M potassium acetate (pH 4.8) was added and the tube was mixed by gentle inversion and then incubated on ice for 5 min. The sample was then centrifuged at high speed for 10 min and the supernatant transferred to a new microfuge tube. One unit of RNase A was added and the tube was incubated for 20 min at 37 C. The solution was extracted twice with 400 ul of chloroform. Two volumes of ethanol were then added to the aqueous phase and the tube was incubated at room temperature for 5 min. The DNA was then collected by centrifugation at high speed for 10 min. The pellet was washed once with 500 ul of 70 % ice-cold ethanol, air dried, and then resuspended in 32 ul of water. Six ul of 5 M NaCl was added, followed by the addition of 38 ul of 15 % polyethylene glycol (8000). The sample was then incubated on ice for

20 min followed by centrifugation at high speed for 15 min. The pellet was then washed with 500 ul of 70 % ethanol, air-dried and resuspended in 20 ul of TE.

#### **Week 5: Quantification of Plasmid DNA**

Purified plasmid was quantified using a fluorometer and a spectrophotometer by measuring absorbency at 260 nm. In addition a fluorometer was used for quantification. Comparison of these two methods was done to familiarize students with different methods of nucleic acid quantification. Based on DNA amounts isolated at this point in the course some students were required to re-isolated nucleic acid from their plasmids due to low yields.

#### **Week 6: Cycle Sequencing**

Double-stranded plasmid DNA was sequenced using an ABI PRISM™ Dye terminator sequencing reaction kit (Perkin-Elmer). Reactions were initiated with 3 pmoles of either the T3 or the T7 primer and were composed of 4.0 ul Dye Terminator Ready Reaction mix, 4.0 ul of HALF-TERM (Genpak, Inc), 500 ng plasmid DNA in a final reaction volume of 20 ul. Reactions were incubated according to the manufacturer's instructions utilizing a Perkin Elmer 2400 PCR system. Thermal cycling was started by rapid ramp to 96 C followed by 25 cycles composed of 96 C for 10 seconds, 50 C for 5 seconds, and 60 C for 4 min. Reactions were then held at 4 C until processing. Free dye terminators were removed from the sequencing reactions by centrifugation of the sample over a CentriFlex Column (Advanced Genetic Technologies Corp) according to the manufacturer's instructions. The column flow through was dried down using a vacuum centrifuge. Sequence reactions were processed by the Nucleic Acid Sequencing Facility (Penn State University, State College, PA). Sequencing gels were not performed in the student laboratory because of cost and because the new fluorescent technology will most likely be remove as a standard procedure in many molecular biology laboratories.

#### **Week 7: PCR analysis of gene inserts**

Each student determined the size of the insert within the plasmid they isolated by amplification using polymerase chain reaction (PCR) followed by agarose gel analysis. The primers for amplification were the T3 and T7 regions flanking the inserted potato cDNA. The reaction was established by placing 25 ng of purified plasmid in a PCR reaction tube with 2.5 units of Taq polymerase, 10 mM Tris (pH 8.3), 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.001% gelatin, and 10 pmoles of T3 and T7 primers. Reactions were placed in a thermocycler for 40 cycles composed of 1 min at 94 C, 1 min at 45 C, and 1 min at 72 C.

#### **Week 8: Analysis of PCR Products**

Students prepared 1 % agarose gels in Tris-Acetic acid-EDTA buffer. PCR products from the previous week were separated on the gel and compared to DNA standards of known size. Band products were visualized on a U.V. transilluminator

following staining with either ethidium bromide or SYBR Green. Students used this information to determine the size of the potato cDNA inserted into the plasmid they have isolated.

### **Week 9: Reading the DNA sequence**

Raw DNA sequence was returned from the Nucleic Acid Sequencing Facility as an electronic file. Students brought in a computer disk in order to retrieve a copy of the sequence for their individual cDNAs. In addition to data transfer students were shown how to read their DNA sequence. This included information such as position of plasmid sequence information, quality of each sequence retrieved, important information associated with some sequences such as polyadenylation sites and or signals, and putative start and stop codons. Some students were also required at this time to redo their sequencing reactions because of reaction failures. Failure rates for the reactions were usually about 10 %. Re-quantification of DNA amounts is necessary for all failed sequences since this is the most likely source of failure for the reactions. It may also be prudent at this time to process additional clones for sequencing in order to insure that all students have a sequence for the final project paper.

### **WEEKS 10,11 and 12: Web Resources for Sequence Analysis**

This presentation has varied greatly in the last few years since the types and sophistication of software that is available for DNA analysis through the world wide web has increased dramatically. The main resources given to the students are the National Center for Biotechnological Information (URL<sup>1</sup>) and Pedro's Biomolecular Research Tools (URL<sup>2</sup>). During the final 3 weeks of the class each student's project took a unique direction. Most students utilized DNA and deduced amino acid sequence information as the main focus of their projects. This included using the blastn algorithm to determine sequence similarity at the nucleic acid levels and the blastx algorithm to compare deduced amino acid sequences between their isolated clone and the databases at Genbank, EMBL and Swisprot (Altschul et al., 1990; Altschul et al., 1997).

### **WEEK 14: In Class Presentations and the Final Paper**

Students were required to present a ten-minute in-class presentation of their results. A final paper in the form submitted to a refereed journal was also required. This paper contained all results, sequence searches, and literature review pertaining to the cDNA they isolated. In addition students were required to develop a proposal outlining future research investigating their individual cDNA.

### **Class Results**

The following is a list of clones currently isolated from potato. The ones with an asterisk we have some expression data for and have published or are in the process of putting a publication together, probably an abstract for a meeting poster presentation.

CLONE	Sequence Similarity
HOW4	germin (oxalate oxidase)*
STW	EREBP transcription factor*
802C	lamin
ELF1	extensin
CSPUD1	glutaredoxin
ELF	extensin
JIG	HMG CoA Reductase
JLS	Wound-inducible Protein pT52
PM	Wound-inducible Protein WIN2
MKP	DNA Binding SPF1 (Yam)
MLA	Chlorophyll a/b binding Protein
CDO	Pathogenesis Related Protein
NHT	Transport Inhibitor Protein (Arabidopsis)
LEC	Transcription Factor (Bean)
JEH	Extensin-like
TAH	cytochrome c
KAS	dioscorin precursor (YAM)
12 Clones	Unique/Uncharacterized

#### Creation of the Database:

Sequence data and similarity results for each clone are currently being placed into a publicly accessible database using MicroSoft Access. This database contains raw sequence information plus Blast similarity searches, background information about highly similar sequences, and it will include the names of the students involved in the isolation and sequencing of the DNA.

#### Conclusions:

Student response to the course was positive based on a survey conducted at the end of the semester. Many of the students enjoyed the concept that they were investigating a gene sequence that was never isolated from potato before even if a gene of very similar sequence was previously isolated from another species. This developed a sense of project ownership for the students and many were very enthusiastic about their project outcome. Some sequences isolated showed no similarity to any GenBank accessions. These unknown sequences presented the most challenge to the students in the class. It was suggested that students trying to analyze an unknown cDNA should focus on deduced amino acid motifs for potential similarities to proteins of known function.

The requirement for a research proposal by all students was also positively received. Behrend College has a small competitive grants program for undergraduate research and some students used the isolation of their cDNA to launch a senior research program.

## Appendix A-Media and Solutions

### LB Broth (liter)

10 g NaCl  
10 g tryptone  
5 g yeast extract  
pH 7.0

### LB Agar (liter)

LB Broth plus 20 g of agar

### NZY Broth (liter)

5 g NaCl  
2 g MgSO<sub>4</sub>·7H<sub>2</sub>O  
5 g yeast extract  
10 g casein hydrolysate  
pH 7.5

### Terrific Broth (liter)

900 ml water  
12 g tryptone  
24 g yeast extract  
4 ml glycerol  
Autoclave, then add 100 ml of sterile 0.17 M KH<sub>2</sub>PO<sub>4</sub>, and 0.72 M K<sub>2</sub>HPO<sub>4</sub>

### SM Buffer (liter)

5.8 g NaCl  
2.0 g MgSO<sub>4</sub>  
50 ml of 1 M Tris (pH 8.0)  
5 ml 2 % gelatin  
Autoclave

## Appendix B

Product	Source
SYBR Green	Molecular Probes, Inc.
XL1Blue-MRF' Cells	Stratagene Cloning Systems
SOLR Cells	Stratagene Cloning Systems
ExAssist helper phage	Stratagene Cloning Systems
ABI PRISM <sup>™</sup> Dye terminator sequencing reaction kit	Perkin-Elmer

**Preparation of Host Cells:** E coli cells for phage infection are grown overnight in LB broth supplemented with 0.2 % maltose and 10 mM MgSO<sub>4</sub>. The following day the cells

are collected by centrifugation and resuspended in ½ the original volume using ice-cold 10 mM MgSO<sub>4</sub>.

NOTE: Aliquots of this library are available for educational purposes by contacting the author. It is requested that any sequences isolated be deposited in the database currently under construction at Penn State Erie.

**References:**

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990)  
"Basic local alignment search tool." J. Mol. Biol. 215:403-410.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z.,  
Miller, W. & Lipman, D.J. (1997) "Gapped BLAST and PSI-BLAST: a new  
generation of protein database search programs." Nucl. Res. 25:3389-3402.

Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) Molecular Cloning-A Laboratory Manual  
Cold Spring Harbor Press, Plainview, NY

URL<sup>1</sup> <http://www.ncbi.nlm.nih.gov>

URL<sup>2</sup> [http://www.public.iastate.edu/~pedro/research\\_tools.html](http://www.public.iastate.edu/~pedro/research_tools.html)