DEVELOPING LABORATORY RESEARCH TECHNIQUES FOR AN ONGOING RESEARCH PROGRAM IN A HIGH SCHOOL CLASSROOM

By

PHILIP ADORNATO

A thesis submitted to the

Graduate School – New Brunswick

Rutgers, The State University of New Jersey

and

The Graduate School of Biomedical Sciences

University of Medicine and Dentistry of New Jersey

In partial fulfillment of the requirements

For the degree of

Masters of Science

Microbiology and Molecular Genetics

Written under the direction of

Dr. Andrew K. Vershon

And approved by

_____

_____

_____

New Brunswick, New Jersey

October, 2010

ABSTRACT OF THE THESIS

DEVELOPING LABORATORY RESEARCH TECHNIQUES FOR AN ONGOING RESEARCH PROGRAM IN A HIGH SCHOOL CLASSROOM

By PHILIP ADORNATO

Thesis Director:

Dr. Andrew K. Vershon

Incorporating research into a high school classroom is an excellent way to teach students fundamental concepts in science. One program that incorporates this approach is the Waksman Student Scholar Program (WSSP), which allows high school students, teachers and Rutgers professors to work side by side on an ongoing molecular biology research program. Students in the program first isolated plasmid clones from bacteria that contain cDNA fragments of genes from the Brine Shrimp *Artemia franciscana*. They then determined the size of the DNA by performing molecular biology experiments. Students then analyzed the DNA sequence and after review from WSSP staff and high school teachers, the student's sequences were published in the National Center for Biotechnology Information (NCBI) database. This was often the last step in the project the students performed. However, if the project were being conducted in a research lab

instead of a high school, the cDNA clone would often be further analyzed. In the past, safety, convenience, and affordability have limited the availability of these experiments in a high school setting. Although additional bioinformatic experiments could easily be performed in the high school, there is a strong need for additional "wet lab" experiments to keep the students engaged and motivated to work on the project. I have worked on developing three experimental modules that can be performed in a high school setting. These experiments were tested with the students and teachers of the WSSP. This work will expand the scope of experiments that can be performed in a high school environment.

ACKNOWLEDGEMENTS

I would like to thank Dr. Andrew Vershon for giving me the opportunity to work in his lab and guide me through my research.

I would like to thank all the past and present members of theWaksman Student Scholar Program for all their help, especially Sue Colleta for her patience and guidance.

I would like to thank everyone in the Dr. Vershon's Lab for their patience with me, and teaching me many different techniques and laboratory skills.

I would like to thank all my professors from graduate school for their guidance and love of science.

Lastly, I would like to thank all my friends and family for their support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

I.1 General Introduction

The United States was at the forefront of scientific research and discoveries for the last half of the 21st century. However, a recent study shows that China may soon surpass the United States in science and technology development if the United States continues to trail behind in science and technology education. This is in part because China places great emphasis on educating and preparing scientists and engineers (15). The United States has lagged in this area, and recent speeches given by President Obama talk about how the United States has been losing ground and lagging in the forefront of research and technology (18). He stressed the need for more emphasis on education and training of young Americans to become leaders in science and technology. Providing more funding may facilitate the resurgence of science research by the United States, but captivating the interest of the youth of America requires more than just monetary compensation.

Many students are put off by science from an early age. One of the main reasons for this is the way science is commonly taught in schools. Memorizing facts, copying of notes, and listening to teachers lecture for entire periods causes students to lose interest. Teachers need to make science more enjoyable and exciting. Teaching current topics and having the students participate in labs and research projects have been shown to improve their overall understanding of science (5).

To compete with the rest of the world and produce new scientists that can make important discoveries and better human lives, high school students should be exposed to genuine research projects that may spark their interest and motivate them to consider a

career in science. Making research techniques and lab equipment accessible to high school students allows them to conduct genuine research at their school during the course of an academic year or longer and exposes them to a world they may have never known existed.

Not all students involved in these practices will become scientists. However, the goal of science education should give students the ability to appreciate and understand the process of science, as well as to think critically of the world around them. These students will become adults and parents, voting on key issues that could affect billions of people in the world. Exposing them to scientific research will allow them to make more informed decisions based on truths and facts, not just hearsay and myths.

I.2 Science Inquiry

In a majority of high schools around the country, science is taught like most other subjects that include notes from lectures and tests, but science also includes a laboratory portion. Most high school laboratories involve experiments with known outcomes that students either know before hand or are directly led to (5). "Schools do not reflect the core attributes of authentic scientific reasoning. The cognitive processes needed to succeed at many school tasks are often qualitatively different from the cognitive processes needed to engage in real scientific research" (5). One aspect is that inquiry based learning may lead to different and unexpected outcomes, which usher students on journeys in which they have to make inferences and new discoveries. A second aspect is that an authentic research project allows students to experience science as a scientist and utilize equipment and techniques that they would not likely come across in a standard

high school lab. Establishing an actual research-based project allows high school students to be engaged in conceptual thinking, problem solving, real world connections and experiences, and linkage of multiple topics throughout a course. "Students at all grade levels and in every domain of science should have the opportunity to use scientific inquiry and develop the ability to think and act in ways associated with inquiry, including asking questions, planning and conducting investigations, using appropriate tools and techniques to gather data, thinking critically and logically about relationships between evidence and explanations, construction and analyzing alternative explanations, and communicating scientific arguments" (7). Not only would scientific inquiry provide students with an advantage if they wanted to pursue research in college and beyond, but it also promotes a greater understanding on scientific concepts and makes students more scientifically literate. By giving them intellectual tools to make import decisions throughout their lives, they are on their way to becoming functional citizens in society.

It is not just important to come up with a research program that allows students to perform research like university and industry scientists, but the project also needs to have meaning to the students, and be based on a topic that they can understand and have a desire to be a part of. Promoting the interests of students helps facilitate any learning process. Because our society is so technologically driven, there is a plethora of topics to choose from that effect the student's lives and ignite their curiosity.

I.3 National Science Education Standards

According to the National Science Education Standards, the emphasis on science educations needs to shift from the standard methods of memorizing facts, in which

students cover an abundance of topics towards fewer concepts (7). There is also a tendency to separate science knowledge from the scientific process. Teachers should have students study subject matter disciplines for their own sake, and implementing inquiry as a set of processes (7). The new method that the National Science Education Standards would like to promote combines some memorization of basic facts to build a foundation, but adds to it with methods that have been researched in the science and educational fields. The new standards focus on understanding scientific concepts and developing abilities of inquiry, learning subject matter disciplines in the context of inquiry, technology, science in personal and social perspectives, and history and nature of science (7).

I.4 Results From Published Models of Science-Based Inquiry

There have been numerous studies suggesting that inquiry-based education is superior to traditional education, which in most cases include lectures and lab activities that have preset outcomes (5,6,7,8,9). However, it becomes difficult to measure how much better this form of education is compared to the traditional methods. One method of quantifying the progress of students is to administer pre and post tests, while another involves journal keeping. The overall theme suggests that students have grown intellectually, which makes this type of learning a positive addition to science education in high school. “We assessed student learning about conceptual knowledge of molecular genetics as well as their beliefs about the nature of science. Student conceptual knowledge increased and their beliefs about the nature of science changed to a more tentative perspective” (8). "We examined student learning qualitatively through their

journals, which showed that some students were developing more sophisticated ways of thinking about the issues that were raised in their seminars and laboratory research. These ways included an increased ability to generate hypotheses, consider alternative hypotheses, implement models and logical argumentation in explanations, connect ideas, extend concepts, and ask questions. These results suggest that meaningfully engaging pre-college students in the practice of real science can make a difference in their understanding and beliefs" (8). Students tend to acquire a greater conceptual knowledge in science when they are emerged in inquiry-based learning. Surveys of students before and after being exposed to inquiry based leaning, showed a positive change in the understanding of concepts, and the students that participated showed gains in all categories (12).

Not all inquiry-based science education works perfectly and there are some drawbacks to this method of educating for the teachers, students, and school districts. First is that the teachers need to invest a copious amount of time for preparation and assessment. They also need to make sure they have an adequate science background, as well as the confidence to administer the research project successfully. Time also plays an import role, because classes are usually under 50 minutes and experiments can be unpredictable and often last more than the allotted time. Teachers therefore need to know when students can stop and continue the next time class meets. Inquiry-based learning will often require additional lab equipments and materials that may be difficult to obtain within a school's budget. Keeping students on task will require the teacher to be more proactive and attentive to make sure students are following directions. Finally, not all students will have uniform experiences. Because these labs are discovery based, students

may draw different conclusions or be led in different directions, while still being correct. This causes a problem with multiple choice styled tests. Tests therefore have to be designed that are more open ended with reflective questioning. However, these types of questions are more difficult to design and time consuming to grade.

I.5 Experimental Modules

For a successful research program in a high school environment to succeed, lesson plans and detailed instructions must be laid out in an easy to read and a well-organized format. Most high school classes are under 50 minutes and may not meet every day of the week. Laboratory modules must therefore fit within the time constraints of school period. In addition, constraints of space and equipment must be taken into consideration. This thesis describes three separate experimental modules that can be integrated in an ongoing high school research program. These modules can be incorporated into genetics or molecular biology curriculum of a class or club, and can be carried out over several weeks.

The concepts for these experimental modules were derived from the WSSP, where students worked on an ongoing research program in collaboration with university research scientists, graduate students, and high school teachers. The WSSP project incorporated both bench top lab work and computer-based analysis. Students started the project by plating bacteria that were transformed with an *Artemia franciscana* cDNA plasmid library. They then picked colonies of bacteria and placed them in a liquid media to allow the bacteria to grow overnight. The next day students harvested and lysed the bacteria and isolated the plasmid DNA in a process known as a miniprep. Once students

isolated the plasmid DNA, they digested it using restriction enzymes. They would also set up polymerase chain reactions (PCR) to amplify the cDNA insert on the plasmid. The PCR products and restriction digests would then be loaded onto an electrophoresis gel to determine if there was an *Artemia* cDNA insert on the plasmid and how large the inserted piece of DNA is. Plasmids containing cDNA inserts larger than 400 base pairs long were then sent to be sequenced. The students then analyzed the sequences using a computer program developed by professors at Rutgers University called DNA Sequence Analysis Program (DSAP). The DSAP program is used to guide the students through this process and to store their answers during each step in the analysis. The DSAP is also used by the WSSP to review the students' data and progress and provide feedback to the students. Students used the National Center for Bioinformatics (NCBI), an online program, to find similar matches to their sequences. Using a suite of BLAST programs, the students looked for similar matches in the DNA and protein sequence database. The students then predicted the protein sequence coded by the protein and performed literature searches to determine the function of the protein in the cell and organism. Finally, the students submitted their sequence analysis for both professors and teachers to verify that the analysis was done correctly. If everything was correct, the sequences were submitted for publication on the NCBI website. The students would then move on and repeat the analysis on another DNA sequence.

At both the end of the WSSP summer session and the end of the academic school year, students present their findings at a poster session in front of Rutgers professors, graduate and undergraduate students, scientists from industry, high school teachers and fellow high school students. Their work is critiqued and discussed. This gives students a

sense of achievement and a feeling of acceptance by the scientific community.

If a similar project was carried out in academic or in an industry research environment many additional steps may be taken to the further the research project. For example, if the sequence of the DNA insert is incomplete (which will be discussed later in this thesis), scientists would try to acquire the entire sequence using primer walking or subcloning. Scientists may also try to clone the gene from a genomic library to identify the positions of introns that were removed from the mRNA that was used to construct a cDNA library. They also would look for the levels of expression of the gene and during what periods in the organism's life the genes are expressed. In the past, these experiments would require sophisticated techniques such as construction of a genomic library, Southern blots, and screening the library through colony hybridization blots. These techniques would require radioactive probes to hybridize with DNA fragments on the nitrocellulose blots. This analysis would be prohibitively expensive, dangerous for students, and involve sophisticated techniques. In addition, these experiments may take months for a fulltime research scientist in a well-equipped lab to complete. These approaches are therefore not feasible in a high school setting. However, the combination of modern laboratory methods, along with decreased costs for some reagents, and the presence of more sophisticated equipment in the high schools, will now enable high school students to perform subcloning, genomic cloning and determine gene expression. This thesis describes the development of some of these experimental modules.

Besides furthering the research, maintaining student's interests and excitement throughout the project is of the utmost importance. Data collected during 2008 and 2009 WSSP project by evaluators from WestEd stated that "students, in large part, did not

enjoy using the DSAP as much as they had the wet labs" (19). From personal experience during my five years participating in the WSSP, a majority of students have expressed their desire to perform more wet lab research because it was more exciting than the bioinformatics analysis. Data collected from WSSP student surveys also shows this (Table 1). This point is further displayed by the analysis of the student's work over the last 3 years. During this period students purified and sequenced over 3000 clones. However, only 15% of the sequences have been analyzed and submitted to NCBI. This data shows that students can easily perform the wet lab experiments of isolating and analyzing plasmid DNA and sending it out to be sequenced. However their bioinformatic analysis lags way behind and as a result student's often never analyzed many of their sequences.

This thesis focuses on furthering the student's research and maintaining their overall enjoyment and interest throughout the research project. Our theory is that placing additional wet lab experiences after the bioinformatics analysis, students will have an increased motivation to work through the bioinformatic portion so they can perform additional wet lab experiments. These experiments will build upon bioinformatic analyses of the DNA sequences (Figure 1). The students will need to revisit and reevaluate their sequences using bioinformatics after the additional wet lab experiments are performed, which will reinforce many of the concepts they learned when performing the first round of bioinformatics. These types of experimental projects will also allow a continuity of subject material that can relate multiple topics within a curriculum, instead of performing unrelated experiments just for the sake of teaching students techniques and topics. The three main foci of the experimental modules are:

1. If there is a gap between the forward and reverse sequences, use plasmid subcloning to determine the sequence gap.

2. Clone the genomic region corresponding to the cDNA

    a. determine if it has introns?

    b. determine if there are polymorphisms?

3. Analyze gene expression at different time points of development.

Table 1: WSSP Summer Institute Student Survey

Year	Percent of students that prefer laboratory work over computer work
2006	80%
2007	86%
2008	94%
2009	77%

The information above was taken from surveys given to the students at the end of the WSSP summer Institutes.

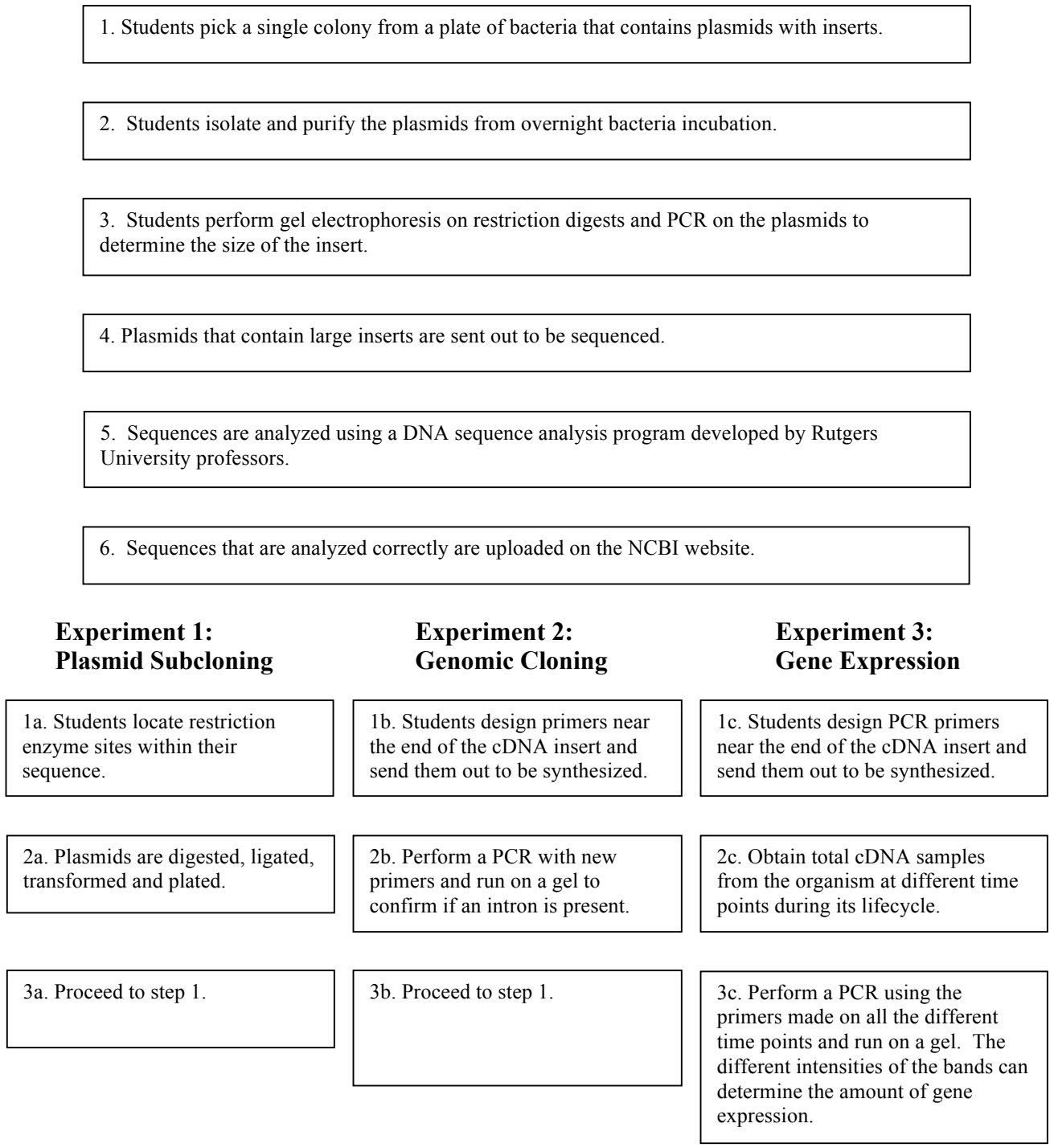


Figure 1: This flowchart represents the existing labs for the WSSP program along with the three new modules that feedback into the existing program.

Chapter II: Teaching Module I - Subcloning
II.1 Subcloning

Students in the WSSP isolated plasmid DNA from the bacteria and then performed restriction digests and PCR followed by gel electrophoresis to determine the size of the *Artemia* cDNA insert. Clones with inserts over 400 base pairs in length were sent out for sequencing. Forward and reverse sides of the insert were sequenced using the SP and XP primers to obtain complete sequence of the inserted fragment (Figure 2). Students performed a blast sequences alignment to determine if they overlap. If they overlapped, a composite (contig) sequence was generated and then used for further analysis (Figure 2a and 2b). However, if the sequences did not overlap (Figure 2c), then the students were missing a portion of the cDNA insert and were not be able to generate a complete sequence of the clone. In previous years, students analyzed the forward and reverse sequences independently. However, it would be significantly more beneficial to their learning process about DNA and gene structure and to other scientists investigating similar genes if they were able to obtain the complete sequence of the clone. Due to changes in costs and methodologies (that were developed in this thesis) it is now possible for students to obtain these additional sequences. As a first step in obtaining the entire sequence of the clone the students need to determine the size of the sequence gap between the SP and XP primer sequences (Figure 2c). They can do this by comparing their sequence to similar published sequences of the same or similar organisms on NCBI to see how much DNA sequence they are missing. If the missing sequence is less than 200 base pair, students could either resequence the DNA with both primers in hope of obtaining longer sequences, or may be able to generate an overlap by being less conservative in cropping.

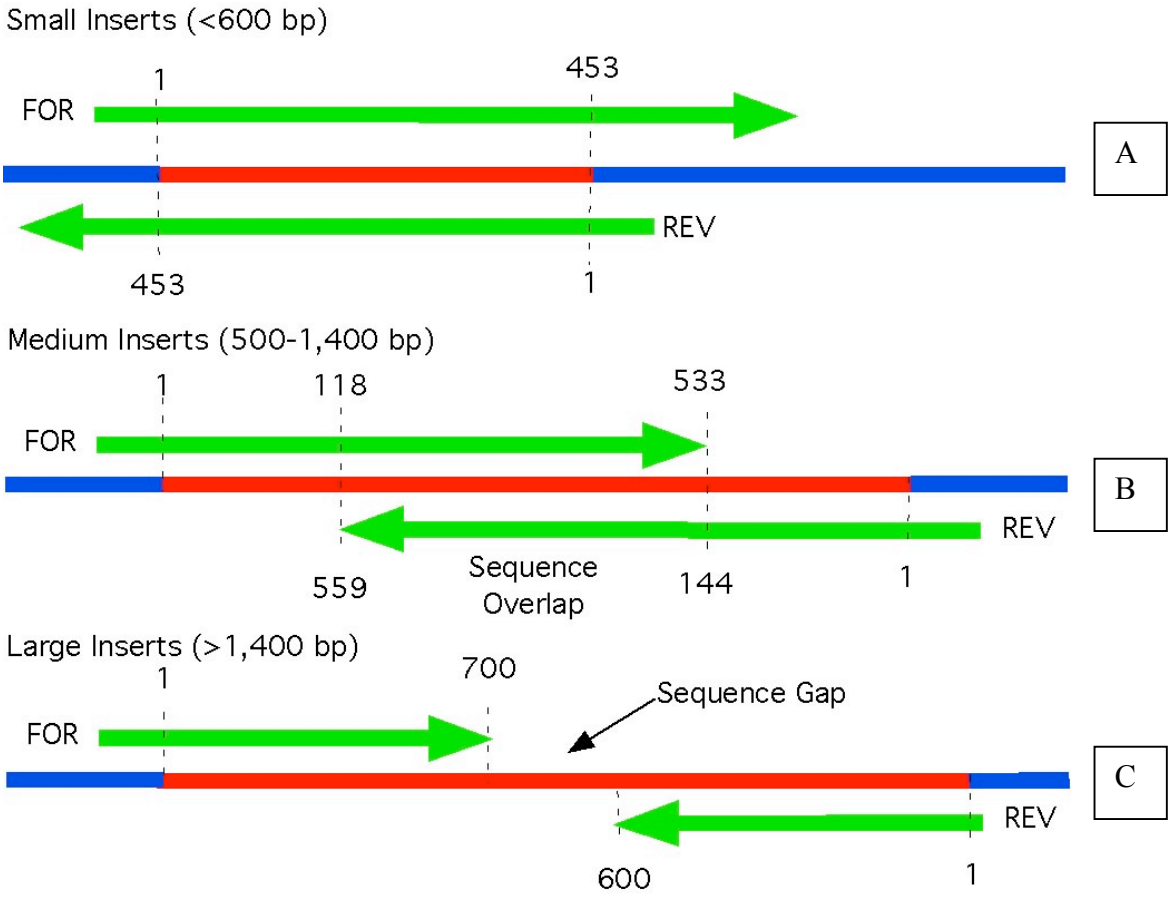


Figure 2: Schematic of possible outcomes in determining if the forward and reverse sequences overlap. The red regions indicate the cDNA insert on the plasmid and the blue regions correspond to the vector on either side of the insert. The green arrows represent the sequences generated by the Forward (For) and Reverse (Rev) primers. These are 3 different situations a student may come across when analyzing their sequence. A) Shows both For and Rev sequences completely overlapping. The Rev sequence can be used to verify the For sequence. B) There is a small overlap between the For and Rev sequences, so that a contig sequence can be formed that spans the entire insert. C) There is a gap between the For and Rev sequences and to obtain the missing fragment subcloning or primer walking may be used.

One method to obtain the missing DNA sequence utilizes a technique called primer walking (Figure 3). Oligonucleotide primers can be designed to hybridize onto the known region of the inserted sequence about one hundred base pairs from the 3' end of the sequence. Primers can be designed using an oligonucleotide calculator, such as http://www.basic.northwestern.edu/biotools/oligocalc.html, to determine the melting points, G-C content, and to see if the primers will self hybridize. Once the primers are designed, the primers are synthesized and the insert is resequenced. The cost of oligonucleotides primers and DNA sequencing have decreased, making this method an efficient process to obtain the missing sequence information. However, a significant disadvantage to this approach is that there is no wet lab component for the students to conduct. Once the students design the primers they just wait for the sequence to come back to perform additional bioinformatics. As a result, while they learn a conceptual process, students do not learn any additional lab techniques, have relatively little problem solving to do, with little abstract thinking, and they are not able to extensively build on previous knowledge. The educational value to primer walking is therefore relatively low.

An alternative method to primer walking is plasmid subcloning. Subcloning is the process of modifying plasmid DNA that has already been isolated. By removing a piece of a known DNA sequence from the clone of interest using restriction enzymes, students can ligate the plasmid back together, transform the ligations, count the colonies, pick colonies, perform minipreps, and perform diagnostic digests and PCR to verify the loss of the DNA fragment (Figure 4). This is very useful for gaps larger than 200 base pairs long. In order to perform a subcloning experiment, students first need to reanalyze their sequences and look for a restriction site within their sequence that is the same as the

restriction sites on the plasmid vector near the insert. Once students find a match, they can then digest their plasmid, remove a portion of their insert, and ligate (Figure 5).

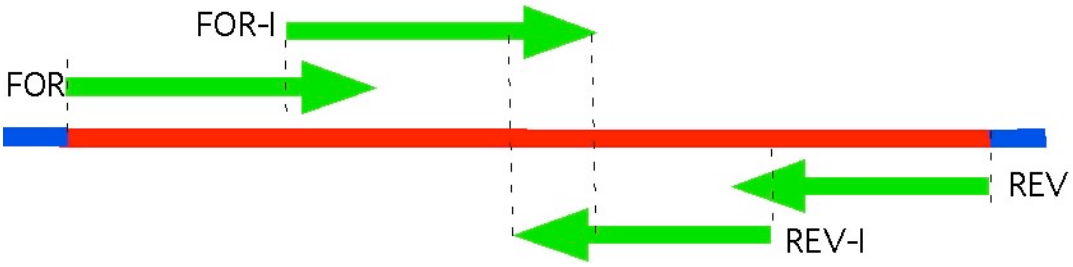


Figure 3: Primer walking can be used to span a small gap between the Forward and Reverse sequences. FOR-1 and REV-1 primers are designed based on the ends of the Forward and Reverse sequences. Sequences generated by these primers may be able to span the gap in the sequence.

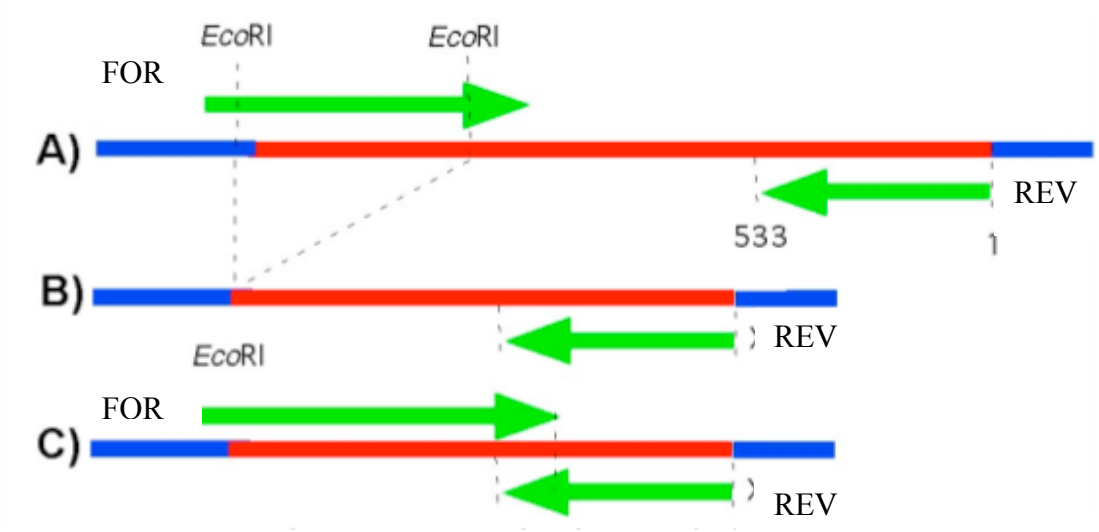


Figure 4: Schematic of Subcloning. Red regions indicate the cDNA insert on the plasmid and blue regions are the regions corresponding to the vector on either side of the insert. Green arrows represent the sequences generated by the Forward and Reverse primers. The position of the *Eco*RI sites are the vector (blue region) and cDNA insert (red region) are shown and can be used to delete a portion of the insert in a plasmid subcloning experiment. A) A clone with a large cDNA insert is shown in which the For and Rev sequences do not completely span the insert. B) Using *Eco*RI restriction digest, a large part of the sequence can be removed and the remaining part of the insert can be ligated back together. C) Once the portion of the insert has been removed, the gapped region can be sequenced with For primer.

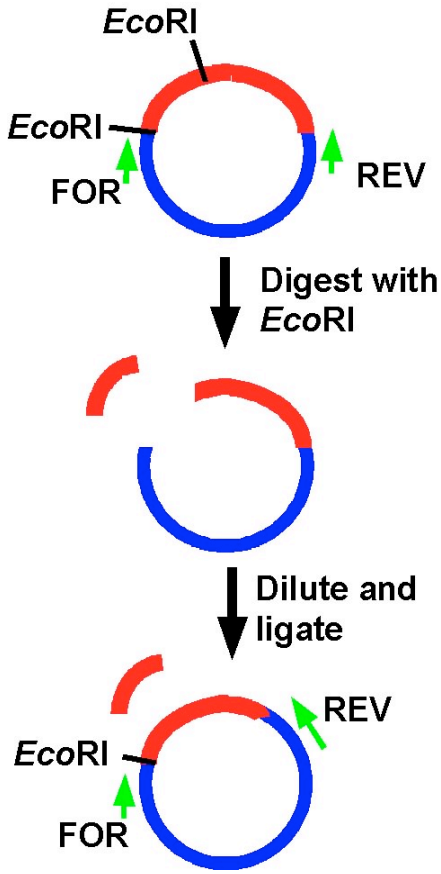


Figure 5: This example uses *Eco*RI as a restriction site within the insert. A portion of the insert is removed performing a restriction digest and the plasmid is then ligated back together. The plasmid will then have to be transformed back into bacteria and students will need to incubate the bacteria overnight in order to obtain colonies for mini preps.

There are many aspects of subcloning that align with the goals of leaning science by inquiry. For example, students learn to think abstractly. There is more than one possible answer that is not predetermined, and these labs are discovery based. Students need to determine the enzyme they will need, the conditions for the enzyme, the sensitivity of enzymes, how to stop the enzyme from completely digesting their plasmid, how to ligate the plasmid, and the use of both positive and negative controls. Since these are discovery based labs and most students have different clones, the answers are not going to be the same and each student needs to derive specific information based on their clone. Students will have to make judgment calls and think conceptually about what they are doing and what would be the best way of obtaining their desired results. A single experimental technique has the ability of teaching students a great deal about molecular genetics. Subcloning is a fairly simple and inexpensive procedure that can be performed in a high school. There are currently products on the market specifically designed for high school student to perform DNA extraction, plasmid isolation, DNA amplification and sequencing, and bioinformatics, but none of them venture into subcloning and utilize the educational potential that it provides.

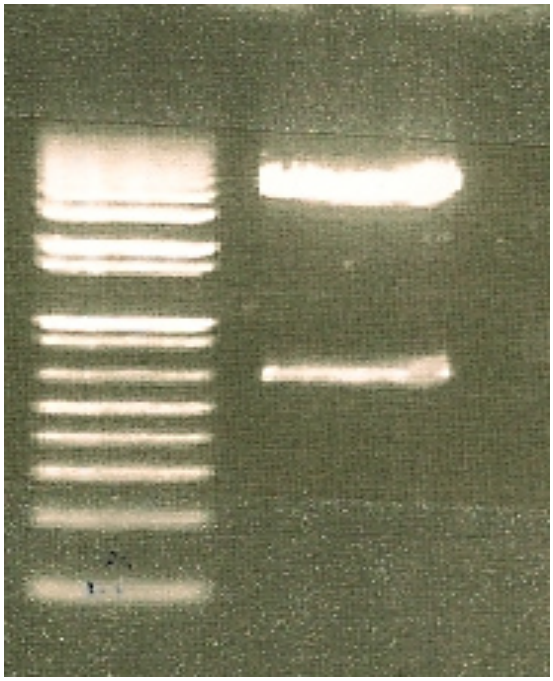
II.2 Methods and Results

The concept and practice of subcloning was introduced to the teachers of the WSSP towards the end of the 2007 summer session. Teachers were given one of four clones in which there was a gap between the forward and reverse sequence, these clones had been chosen by the WSSP staff for the presence of restriction sites that could be used for plasmid subcloning. The entire procedure had limited success. This was easily seen

after the teachers chose their restriction enzyme, digested the plasmid and analyzed the bands on the gel. Many of the digests failed to show a digested fragment from the plasmid. Some teachers did not evaluate their clones correctly and failed to determine the proper restriction sites. This was in part, because the teachers were rushed into this technique with little background and experience. Experimental procedures were compiled at the last minute and lacked optimal control conditions. In addition, during the week that the teachers were performing this exercise they also attended a workshop with entirely different techniques and procedures about protein purification. The information was overwhelming and the professors at Rutgers were not able to provide adequate instruction. All these factors contributed to the limited success in the summer of 2007.

After that summer session, I worked on revising the protocol to optimize the results and overcome problems and obstacles that the teachers had in the procedure and that maybe problematic in a high school setting. For example, clarifying procedures for analyzing the sequence to correctly locate restriction enzymes sites. I took the original procedures that the teachers followed in the summer of 2007 and retried them to test whether they worked or not under more controlled conditions. Clones JL5.07 and 0AV696.06 were used for this analysis. These clones contained two *Eco*RI sites, one in the vector polylinker and one *Eco*RI site in the insert. The plasmids were run on a gel and fragments of the predicted sizes were visible on the gel. This confirmed that the digest was successful (Figure 6). The plasmids were then ligated and transformed and colonies grew and the ligations and transformations were successful (Table 2).

Ladder *Eco*RI digest



---------- Vector backbone

---------- Deleted fragment

Figure 6: This is an *Eco*RI digestion of clone JL5.07 to remove a 650 bp fragment containing a portion of the insert.

Table 2: Transformation of Subcloned Plasmids

Subcloned DNA	Number of colonies	
No Ligase	0	6
Undiluted DNA	Lawn	Lawn
5X Diluted DNA	100	60

Table 2: The results from ligation and transformation of a deletion subcloned plasmid. Transformations were in duplicate. The 5X dilution yielded the best results for students to pick individual colonies for minipreps and sequencing. Subcloning produced enough colonies that the plasmids can be harvested, isolated and the inserts sequenced. Undiluted DNA made it too difficult to pick an individual colony.

In the 2008 WSSP Summer Institute the teachers again performed the subcloning procedure using the revised protocols, equipment, enzymes and chemicals. Teachers were giving different clones to analyze and had to determine which restriction enzymes were needed to delete part of their insert. After the teachers digested their plasmid with the enzyme they verified the digest by running it on a gel to determine if the enzyme removed their desired fragment. They then ligated and transformed the plasmid into bacteria. Colonies were then picked and overnight cultures where grown, and teachers performed minipreps to isolate the plasmid DNA which was verified by PCR and restriction digest. As a result, more than 75% of the teachers were able to perform this technique and confirm their success on a gel, and are experienced enough to introduce and implement subcloning experiments back at their schools.

Most of the steps in the subcloning procedure use reagents and techniques that the teachers are already familiar with. One exception to that is the transformation of the ligation reactions into competent cells. The transformation procedure can be problematic, because the cells that take in the plasmid are extremely sensitive to environmental conditions. These cells are known as competent cells and are specially engineered to take in plasmid DNA. They are stored at -80 ^{0}C and only taken out and defrosted on ice to be immediately used. They will lose their ability to take up plasmids if they are left out too long or refrozen. This is problematic because virtually no high school has a -80^{0} C freezer. In fact, most high schools only contain frost-free consumer models of kitchen refrigerators that automatically cycle above freezing. To test if this is a problem, experiments were carried out to determine if the competent cells that are stored in a typical freezer found in high schools can be transformed with plasmid DNA. Thirty

super competent cells were placed into a Styrofoam microcentrifuge tube rack and placed into a -20^0 C frost-free freezer. Each transformation was done with the plasmid 9PA08.07 isolated from the Waksman Student Scholar Program. The plasmid was transformed using 3 different dilutions, 1:100, 1:1000, and 1:10000. Transformations were done over a one-month period, starting with the first transformation performed after having the super competent cells sit in the -20^0 C for 24 hours. The transformations were performed at 2, 8, 9, 15, 22, and 29 days of being in the -20^0 C freezer. Along with each transformation, a control was used which is the zero time point. The controls were competent cells taken directly from the -80^0 C and put on ice, and transformed along side the plasmids from the -20^0 C cells with the same dilutions of DNA. Results from these experiments show these cells remain viable and competent for up to 3 weeks in a -20 ^{0}C frost-free freezer (Table 3).

Only two high schools in the 2008-2009 WSSP school year attempted to subclone. This is in part because they received the DNA sequences of their clones very late in the school year, and the students did not have enough time to run the experiments. Both schools used NEB cutter and located regions on their insert to be removed and chose restriction enzymes, but never ordered the enzymes and digested the clone. There was increased interest from the high schools to carry out this cloning during the 2009-2010 WSSP school year. A new cDNA library of the duckweed *Wolffia arrhiza* was used during the 2009-2010 WSSP project. Unfortunately most of the clones in this library were shorter than 500 base pairs and therefore the complete sequence of the cDNA insert could be obtained with the forward primer, or if needed, by overlap between the forward and reverse primers. Therefore only a few (less than 10 of the 1000 clones sequenced)

had a sequence gap. Examination of the sequences of these clones did not reveal any useful sites for subcloning.

Table 3: Stability Test of Competent Cells

Dilution of Plasmid	**1:100**		**1:1000**		**1:10000**	
Temperature ^{0}C	**-80**	**-20**	**-80**	**-20**	**-80**	**-20**
Day 1	Lawn	280	380	300	58	14
Day 2	Lawn	Lawn	320	31	34	33
Day 8	n/a	n/a	57	2.5	6	0
Day 9	300	33	13	1	5	1.5
Day 15	174	20	5	0	n/a	n/a
Day 22	200	26	10	0	n/a	n/a
Day 29	150	0.5	0.5	0	n/a	n/a

Table 3: This table shows the period of time super competent cells can be kept in a -20^0 C frost-free freezer. The different concentrations of plasmids were transformed into competent cells stored at -20^0 C and -80^0 C for the indicated number of days. Each colony count represents the average of 2 plates. The plasmids for the transformation were diluted. Therefore, 1:100 dilutions are as good after one week, and maintain 10% viability after 29 days.

II.3 Discussion

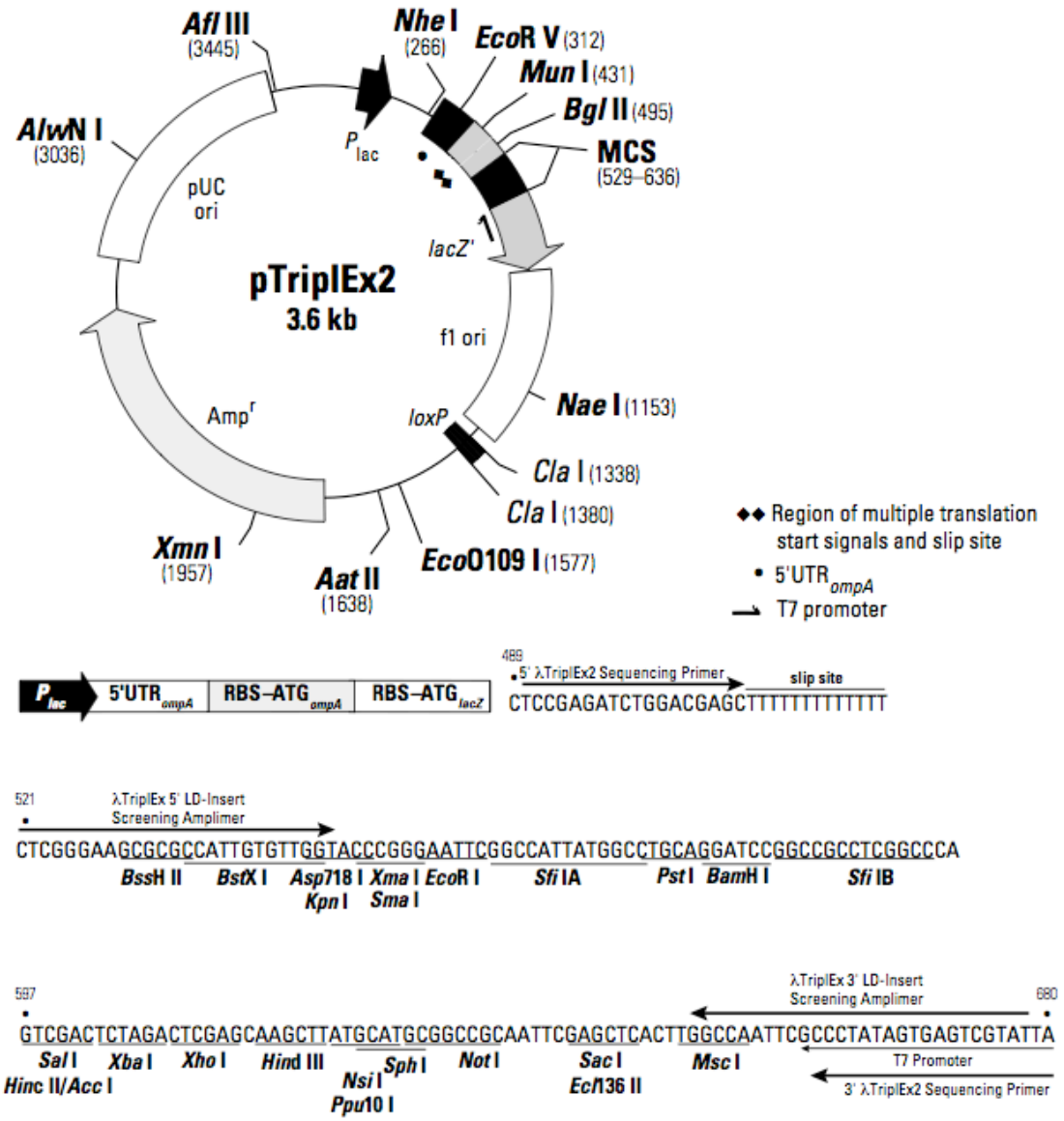
There were many mistakes and some problems that needed to be worked out the first time subcloning was introduced to the teachers in the WSSP. Many teachers never heard of this technique before and in less than a few days they had to learn what it was and perform multiple experiments. Procedures were loosely drawn out and there was much confusion. After the WSSP Summer Institute ended, we reflected and tried to refine the procedures. After many variations of previous experiments, the above procedure worked the best and had extremely high reproducibility.

We found that many of the teachers were new to this technique and may have been overwhelmed the first time they performed the technique, but after trying it a few times it became very simple and teaching it to the students is not a problem. Being careful and meticulous will help insure better results. None of these steps should be rushed. Places in the protocol were identified to stop and continue the next day if the students were behind schedule in the lab period. Teachers will need to provide significant oversight in the process and make sure students pick the right enzymes and follow the procedures correctly.

All of the reagents used for subcloning were supplied by the WSSP. However, if schools want to perform additional analysis on their own, all of the reagents can be purchased. The costs for this module should be within a high school's budget. The most expensive part is the purchase of the restriction enzyme, ligase and competent cells. Teachers should only purchase the enzymes that are required for subcloning (Figure 7). For example, restriction enzymes in the 5' direction *BstX1*, *KpnI*, *SmaI*, and *EcoRI,* and in the 3' direction *XbaI*, *XhoI*, *HindIII*, *NsiI*, *SphI*, *NotI*, *SacI*, and *MscI*. These enzymes

are located on the plasmid vector and flank the inserted piece of DNA. If handled carefully and depending on the usage, these enzymes should last a few years.

If this module is to be incorporated into the class curriculum, a way of assessing the students during class, would be to set up rubric. Important information students should walk away with would include identifying restriction sites and determining the conditions needed to digest the plasmid. They should be able to determine if a digestion occurred and predict what size the bands should be on gel. They should also be able to determine if there were more than one restriction site in the unknown region of the insert using gel electrophoresis. Once they determine if the digestion was successful, the students then need to be able to ligate the plasmid. The ligations must then be transformed into bacteria. Students should understand the value and reasoning behind the positive and negative controls used during ligation and transformation. For example, students need to set up negative controls with no DNA and unligated plasmid (Table 4). In both of these cases there should be no transformation. If something does grow there is either a contamination or no ampicillin on the plate. The positive control, transformation with uncut plasmid, should form a lot of colonies.



Restriction map and multiple cloning site (MCS) of pTriplEx2. Unique restriction sites are bold.

Figure 7. Diagram of the pTriplEx2 plasmid that was used throughout the Waksman Student Scholar Program. The major restriction sites in the MCS region that can be used in the subcloning experiments are shown.

Table 4: Ligation and Transformation Controls

Test Tube Number	Sample	Expected Results	Conclusion
0	No DNA	No Growth	If colonies grew, there was contamination.
1	Unligated DNA	Less than 10 colonies	If more than 10 colonies grew, digestion did not occur.
2	Ligated DNA	Greater than 10 colonies	If no colonies grew, ligation did not work.
3	Ligated DNA 1:10 dilution	Around 10 colonies	If no colonies grew, ligation did not work.
4	Uncut DNA	A lawn of colonies	If no colonies grew, cells could have been not competent

Table 4: Five different reactions are set up with different samples to determine that there were no false positives or false negatives to validate the experiment.

Chapter III: Teaching Module II – Genomic Cloning

III.I Introduction

The WSSP research project involves the isolation, characterization, and sequence analysis of clones from a cDNA library. The advantage of using a cDNA library was that the clones were likely to be in the coding regions of proteins, and this helped the students identify the proteins for their reports. Because introns are spliced out of mature mRNA, coding for a protein sequence may only be a few hundred bases long, while the corresponding regions in the genomic sequence may be a few thousand bases long (Figure 8).

In general, coding regions among closely related organisms are conserved with relatively minor variations. However, because there is no selective pressure, non-coding regions of the genome are usually a lot less conserved between even closely related organisms. A genomic sequence therefore is useful when researchers want to look at subtle discrepancies among different organisms of very closely related species or even individuals within a species. Students can do this by comparing the 5' and 3' untranslated region (UTR) of their cDNA clones among the different organisms. However, if they also have the genomic sequence corresponding to their cDNA clone they can compare the introns and determine the splice points. Students can look at variations within a species noncoding regions, which may not be as highly conserved as the coding region, if they had access to a genomic DNA. Students could also compare organisms from one habitat to organisms found in another habitat.

Construction and screening of a genomic library for a specific gene by DNA hybridization with a cDNA clone is not suitable for a high school. The use of radioactive

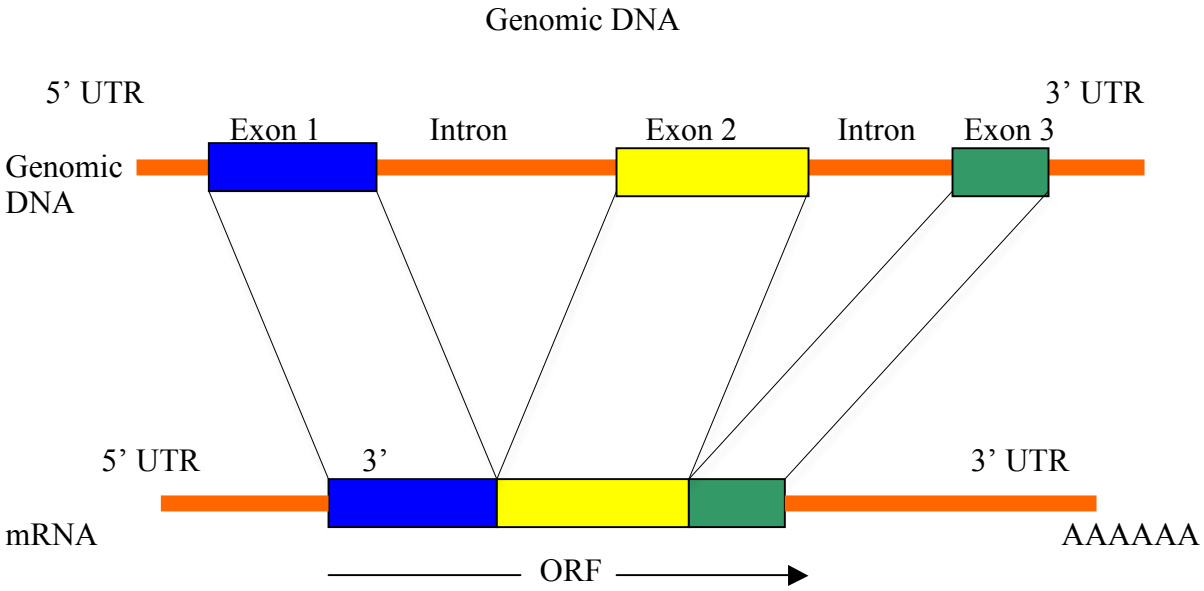


Figure 8: The upper diagram represents a fragment of genomic DNA with introns and exons before transcription. The lower diagram represents a much shorter piece of mRNA after processing. Introns are removed and the exons remain and that code is translated into a protein.

probes, the large effort to colony screen a genomic library, the high quantity of screening, and the low chance for obtaining a sequence makes this method impractical. This approach could take weeks in a university setting with full time researchers. We therefore tested an alternative method to see if a fragment of genomic DNA could be amplified using primers based on the cDNA sequence (Figure 9). Extracting genomic DNA from an organism is relatively simple and can be performed in a few hours in a high school setting. Alternatively, the WSSP staff could prepare the genomic DNA and provide this to the students. Once extracted, DNA is very stable and students can go back to it many times throughout the school year. This module will allow students to see if their clone contains introns, sequence PCR product directly and determine if there are any polymorphisms.

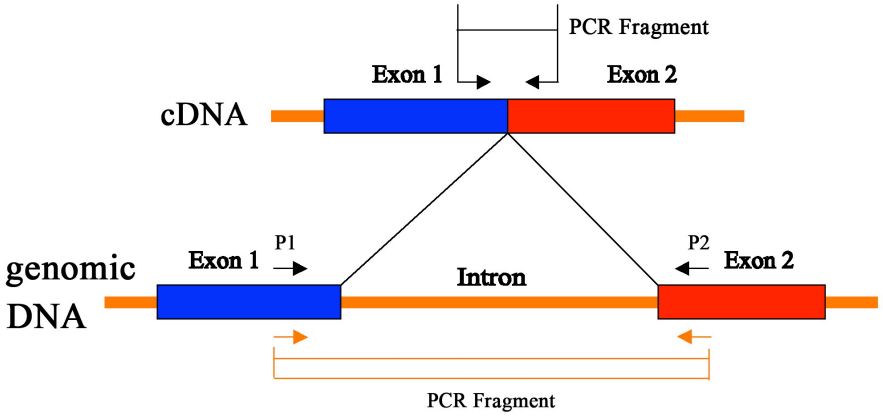
Once students amplify a DNA fragment they believe contains an intron, they will then need to determine the sequence of that intron. Students can add ExoSAP-IT® to the PCR product directly, which cleans up the DNA and can be used directly for sequencing. If the genomic fragment is larger than the cDNA clone, that indicates the presence of an intron.

PCR often adds an additional A to the end of the 3' end of the DNA fragment. This makes cloning into a blunt end of a fragment difficult. However, PCR fragments can be cloned with engineered vector fragments that contain a single T overhang (Fig 9). Companies provide kits, such as TopTA from Invitrogen and pGemT from Promega to aid in this process.

The TopoTA kit is far too expensive for most high school budgets and the amount of reactions per kit were too few to be useful in a class setting. The Promega pGem-T®

vector was far less expensive alternative, and if the transformation was as successful, then this would make a great alternative. We also tested if we could use lower than the recommended volumes to perform the protocol, allowing us to perform a greater number of reactions and further save money.

A.



B.

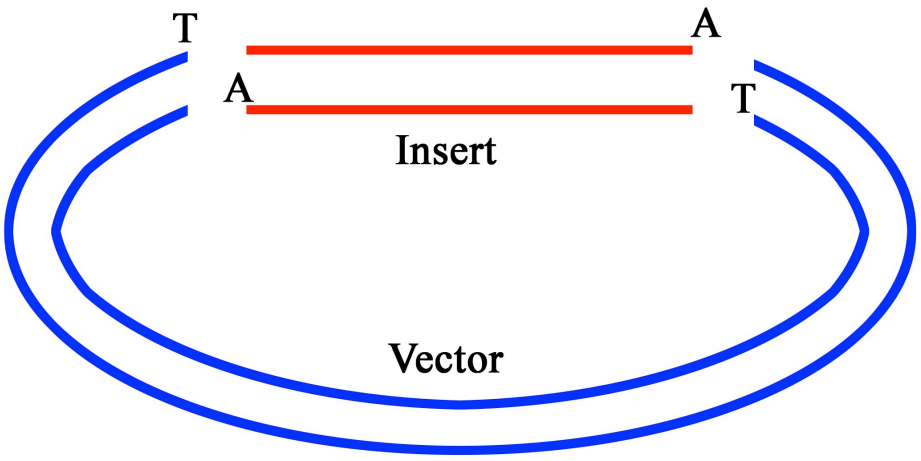


Figure 9: A. Primers were developed near the ends of the two exons that would span the intron based on the genomic sequence in figure 9. B. The vector has T's on the ends to overlap the insert with A's.

III.2 Methods and Results

Artemia genomic DNA was extracted using a self modified "fly squishing buffer" from Dr. Mckim's lab. Adult *Artemia* were taken and spun down into microfuge tubes with a mass of 0.5 grams. 70 µl of this buffer (1M Tris, 0.5M EDTA, 10% SDS, and H_2O) was added and heated for 30 minutes at 70^0 C in a hot water bath. Periodically the sample was smashed up using a sterile plastic pestle. After 30 minutes, the sample was centrifuged on high for 10 minutes and the supernatant was transferred. For every 100 µl of fly squishing buffer, I added 22.5 µl of 5M KOAc and incubated for 30 min on ice. One volume of phenol/$CHCl_3$ was added twice to each sample and the top layer was collected. Two volumes of 100% ethanol was added to the sample and the DNA precipitated out. The sample was spun for 2 minutes and the supernatant was discarded, and washed again with 70% ethanol and spun for 2 more minutes and discarded the supernatant. The sample was dried and 100 µl of water was added and the optical density was calculated.

To test for the ability to amplify genomic DNA I designed primers based on the Hsp26 gene from *Artemia* (accession # DQ310575). The genomic DNA that was amplified from a PCR with the Hsp26 primers was cloned into a *TopoTA*® vector. I aliquoted 4 µl of amplified genomic DNA, 1 µl of salt solution containing 1.2 M NaCL and 0.06 M $MgCl_2$, 1 µl of *TopoTA*® vector and incubated for 5 minutes at room temperature. After 5 minutes, 1 µl and 3 µl of the mixture was added to two separate 100 µl competent cell tubes and incubated for 20 minutes. The reaction was then placed on LB + amp and x-gal plates and left over night. Both white and blue colonies grew and 12 white colonies were picked and incubated in liquid media over night. Overnights were

miniprepped, digested, and ran on a gel. A PCR was also performed using the M13 primers from the *TopoTA*®, producing multiple fragments ranging in size from 700 to 1100 base pairs.

Because of the high price of the *TopoTA*® kit, I set up an experiment to compare the effectiveness of both *pGem*® and *TopoTA*®. The included protocols for each vector's ligation were followed. The only variation was the amount of PCR product used, which I used both 1ul and 3µl PCR product for each. A negative control was used that lacked a PCR product. After one hour of incubation, the samples were transformed into competent cells. Both vectors had similar positive results from the transformation experiments (Table 5 & 6); therefore a cheaper method can be used, because a school budget needs to be kept in mind. The major limiting factor in ligation and transformation is the vector. If the vector can be diluted, schools will be able to perform many additional transformations with the same kit. Another experiment was set up with a 1:10 dilution as well as a 1:20 dilution of the vector. The same procedures were followed as above and incubation time was 1 hour as well as an addition 24-hour incubation time. The results of this experiment produced enough viable colonies using a 1-hour incubation, so the 24-hour incubation is not necessary (Table 5).

Table 5: Results from TopoTA vs. pGEM-T Vector

Vector	1 µl of PCR Product		3 µl of PCR Product	
	# Blue Colonies	# White Colonies	# Blue Colonies	# White Colonies
Topo-TA	688	800	832	960
pGEM-T	320	160	960	640

Table 5: This table shows the difference between the two vectors used for ligation and transformations. Four ligations and two transformations for each were performed, and the average of the white and blue colonies were calculated The pGEM-T is significantly less expensive than the Topo-TA vector and had similar results, which saves money.

Table 6: pGEM-T Vector Dilution

	1 Hour Incubation		24 Hour Incubation	
Dilution	**# Blue colonies**	**# White colonies**	**# Blue colonies**	**# White colonies**
1:10	168	40	24	128
1:20	92	24	80	32

Table 6: This table shows the pGEM-T vector diluted. The ligations were left to sit on ice for either 1 hour or 24 hours and then transferred into the competent cells. This dilution can save schools and the WSSP thousands of dollars. The pGEM-T vector was diluted to see if many more reactions can be used, which could save thousands of dollars. When diluted 20-fold, 80 blue colonies were successfully grown.

To develop the experimental module for use in a high school setting, oligonucleotides were designed from a published genomic sequence of *Artemia* that was in the NCBI database (accession # DQ310575). Primers were chosen that flanked an intron in the genomic sequence of the *hsp26* gene (Table 7). These primers were around twenty bases long and located in the coding region of the gene and flanked a known intron on both sides (Figure 10). The specific primer sequences were chosen to contain a high G-C content of 50% or better with a melting point around 60 ^{0}C. The primers were checked to make sure the sequence is not self complementary to prevent the formation of hairpin loops. To design primers, sequences between 19 and 25 bases long were analyzed in the Oligonucleotide Calculator. (http://www.basic.northwestern.edu/biotools/oligocalc.html). The primer sets were used to amplify fragments using total genomic DNA as a template. The primers were also used to amplify DNA from cDNA pools that were synthesized from mRNA isolated from *Artemia* cysts at 0 and 24 hours after incubating them in hatching media.

Based on the published genomic sequence, the amplified fragment should have been around 1500 bases long. A majority of the sequences amplified from a PCR reaction were very short, around 300 and 850 base pairs long in lanes 2, 3, and 4 (Figure 11). To determine what these fragments were, the PCR fragments were ligated into the Topo-TA vector, transformed into bacteria to yield between 20 and 50 white colonies. Twelve colonies were picked to isolate the plasmid DNA and sequenced. The sequences did not have good matches on NCBI at the time. However, new sequences have been published on NCBI and there are some matches to heat shock protein introns and other nucleotide sequences from *Artemia franciscana* as well as nucleotide sequences from

other *Artemia* species. Some of the sequences matched ends of *hsp26*, which were my oligo's and others matched introns from a zinc finger gene of *Artemia*. The primers did match up on both ends as well as the intron from another species of *Artemia* (Figure 12).

The large number of bands generated in the PCR and the improper sizes of the bands suggested that the primers were annealing and amplifying other regions of the genome instead of the *hsp26* gene. Nested primer sets were therefore used in an attempt to amplify the desired sequence from the PCR. Nested primers are a second set of oligonucleotides that are adjacent to the original primers. The amplification of the DNA fragments with this second set of primers using the amplified DNA from the first primer set should increase the specificity (Figure 13). A total of four primers were used in different combinations in a PCR. Once the PCR was done, the samples were ran on a gel and the DNA fragments were analyzed (Figure 14). A PCR sample was taken from the 1:100 dilution of both outer primers, outer and inner primers and inner and out primers and amplified a second time with just the inner primer set (Figure 15). Because the sizes of the bands were not near the size of the desired intron, those sequences were not cloned or sequenced.

Table 7: **Oligonucleotides**

Gene Name	**Accession #**	**Oligo Sequence**
Heat shock protein 26	DQ310575	5’ outer: GCT ACT CCT GGG TCC TTG AGG G
		3’ outer: CGC TGG TGT GAT GGG TAC GAT ACG
		5’ inner: CCG ACG ATG ATA TTC TTG TCC ATG GC
		3’ inner: GAG CTT AAA GCA GTT TTC GGA GCA TG

Table 7: Oligonucleotides designed from a published sequence of *Hsp26* (Accession #: DQ310575) from NCBI with the goal of cloning out the *Hsp26* gene from genomic DNA.

```
6661 gccaaaacag aatattagaa atggttatat ttgcaggccc ttccggagaa gaatgatgag
6721 aagaggtcca gataccagca gggctttaaa ggagttagct actcctgggt ccttgaggga
6781 cacagctgat gaatttcaag ttcagctaga tgttggccac tttttaccaa acgaaattac
6841 agtcaagaca accgacgatg atattcttgt ccatggcaaa catgacgagc gatctgatga
6901 atatggacac gtccaaagag agtgaggctt atctttttc tcctttcttt ttttttcttt
6961 taatacttag ctgtaggatc ggggatctag agggggaatc gggatgaagc ggagtggagg
7021 aatcgggatg aagcggagtg gaggaatcgg gatgaagcgg agtggaggaa tcgggatgaa
7081 gcttggtgga tggaatatat aaatattcat tatgtccatt aaaatgtcca tgacagcgaa
7141 ttttctgtct ccttactcct gggatattct ccggtgcaat tatgactact tttgcagtct
7201 ccatttttgg acggtaagca catgaatgta atcttttaat gctaaaaaaa gttaaagaac
7261 ttgtctccca gcaacaccca tattaggctt agtgcaaatg aagctaaaat aagacattaa
7321 tttcaactca tctgacacgt tatgggattc cctgctcctt ccacttttat caaatagttc
7381 gtggtaagga actagtaagg agcggccagg cccaaaagta accgacacta aaaaacaaaa
7441 ttttgatacc catagttaca tcaaaagata aacattttaa ttctgatttt gactcttgac
7501 tccacttttt aaaatagtaa aaaatttagc gtaaagagcg gggtgttgag gagggaacaa
7561 ccccttcgt atacggagta atttctgttc gttttaagct ttaatgtcgc tccttacttt
7621 tggttgaaaa aacttaaggt cagttaaaaa actttaattt aacaaagtaa ggatccgaat
7681 gcaacctttt atatcataga atatcaacat ccatattgcg ctcttcttga aaacaactcc
7741 tccaaataac gggaggttag gggtaaatct gtcttgctat gtgggaccgg ctgactatgt
7801 ccatattgat atgtaggaca ggctgactaa tttggcccat attgtgaggc ttccggtcat
7861 tcacggaatg atctgcgtcc tccgacactg aaatatatca ttcttaagat aggaaaactg
7921 atttggatat aaaaatatcg aatttccaac atcttgttaa gcaaagcaga gaacctctat
7981 gaaacactta ggaagattaa gcttaaactt gttaaattta ccccctcaca ccatttaaag
8041 aaaaataagt aataagaatt aaaataggtg caacatatgt ttggcttgtg gattatttgg
8101 cttgtagcca acgcgaatgg atcatggtga tgacttatgc taaaaaattc atgactgttg
8161 gtccaagccc tacttttta catttattat acattatttt gcaggatttcg acgacgatac
8221 agactcccag aacatgtcaa accagaatct gtgtcatcta ctttgtcatc agatggtgtc
8281 ttaactatcc atgctccgaa aactgcttta agctcaccaa cagaacgtat cgtacccatc
8341 acaccagcgc cagctgttgg aaggattgaa gggggaacta caggtactac tacaggcagt
8401 acagctagtt caactccagc aagaacaaca agatcaggag gtgcagctta atctgcattc
8461 aatatatatt caatgaatct tccggtcttt tttctttgta cctttatttt tttgtcaata
8521 aatctgtata cgaggcaatt gaagattgtt cccttcgctt accaaacaat attctcttta
```

Figure 10: This is a partial genomic sequence of the *Hsp26* (Accession #: DQ310575) gene which includes an intron from base pair 6922 to base pair 8204. Primers both inner and outer primers were designed in the exon region of the genomic DNA. Blue represents inner primers, the red represents outer primers, and between the yellow "g's" represent the intron

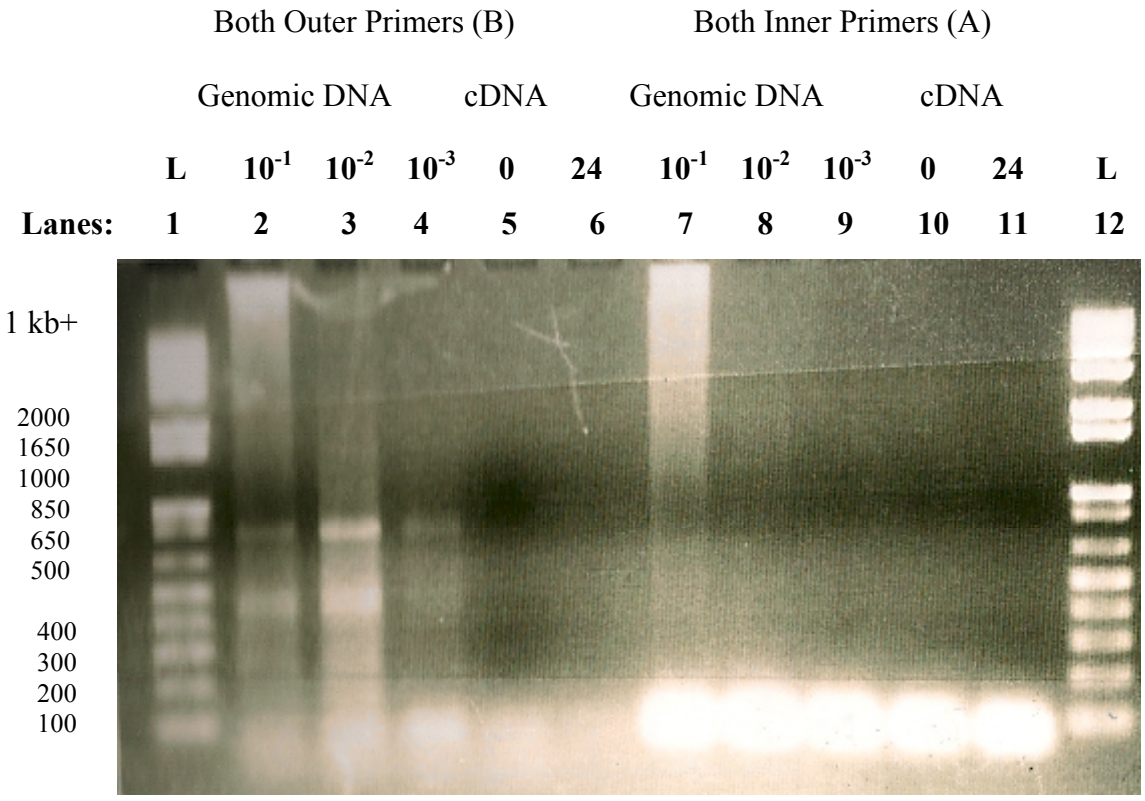


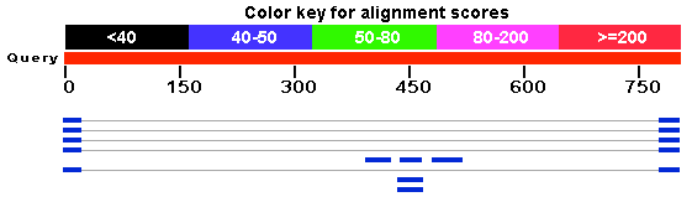
Figure 11: Lanes 2-6 used outer primers (highlighted red from figure 10) in an attempt to amplify the heat shock gene from both genomic DNA and cDNA. Inner primers (highlighted blue from figure 10) are in lanes 7-11were developed and again the heat shock protein was attempted to be amplified using PCR. The outer primers appear to amplify numerous genomic sequences, while the inner primers did not amplify anything. "L" represents a 1 Kb+ ladder, and the numbers below each the genomic DNA represent the dilution of genomic DNA. The numbers below the cDNA represent cDNA that was extracted at 0 hours and 24 hours.

A) **Clone PPA 100 SP vs hsp26**

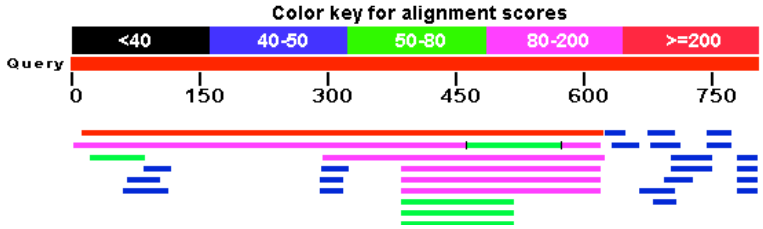
```
Query  8325  ACGTATCGTACCCATCACACCAGCG
8349
             |||||||||||||||||||||||||
Sbjct  26    ACGTATCGTACCCATCACACCAGCG  2

Query  8326  CGTATCGTACCCATCACACCAGCG  8349
             |||||||||||||||||||||||||
Sbjct  656   CGTATCGTACCCATCACACCAGCG  679
```

B)



C)



D)

```
>gb|GU248389.1|  Artemia franciscana isolate CEJ5 p26 gene, intron 2 and partial
cds
Length=1792

 Score =  134 bits (148),  Expect = 6e-28
 Identities = 187/249 (75%), Gaps = 21/249 (8%)
 Strand=Plus/Plus

Query  390   TTAAACCATT---TCACTTTTTAGGGGTGTTTCCCCCT-TTATCAAAAAATTAGACAAAT  445
             |||||  ||||    | ||||||  |||||||||||||| | || || |||||   |  |||||
Sbjct  1309  TTAAAACATTCTATTACTTTTAGGGGGTGTTTCCCCTTATTTTCTAAAAA--ACGCAAAT  1366

Query  446   TTTCTCAGGCTCGA-AACTTTGG-TGGGTAG--CTTATAAC---ATGAA--TTCTATATT  496
             |||||||  ||||   ||||||  | ||||||   | || |||   |||||  ||  ||||||
Sbjct  1367  TTTCTCAGACTCGGTAACTTTTGATGGGTAAGACCTA-AACTTGATGAAACTTATATATT  1425

Query  497   TTGAATAAACATCAAAATTAAATTCTTTTGATGTATCTATTATTATC--AAGTCCCTGAT  554
             |   |||  |  |||  |||||    |||||||||||||| |||||  ||||   ||  |||  |  |
Sbjct  1426  TAAAATCAGCATTAAAATGCGATTCTTTTGATGTAGCTATTGGTATCAAAATTCCAT--T  1483

Query  555   TCTTATAGTTTAGGTTACTATTAAGCCGCATCGCCCCTTGCTTACAGTCTGTTACCACAA  614
             | ||| |||||  |||| |||||  ||||  |||| | ||||||  |||||||| |
Sbjct  1484  TTTTAGAGTTTTGGTTTCTATTGAGCCGGGTCGCTCCTTAC-TACAGTTCGTTACCACGA  1542

Query  615   ACTATTTGA  623
             |||  |||||
Sbjct  1543  ACTGTTTGA  1551
```

Figure 12: The primers that were made did amplify small regions of DNA, but the heat shock protein intron was not. A. My clone is labeled "Sbjct" and the *hsp26* is labeled "Query" sequenced blasted together in a blast2n. This shows that my primers were good. B. Shows results from a Bastn from my clone pPA100. C. Shows another one of my clones pPA004 and the blastn results. D. Shows a match to an *Artemia* heat shock protein intron.

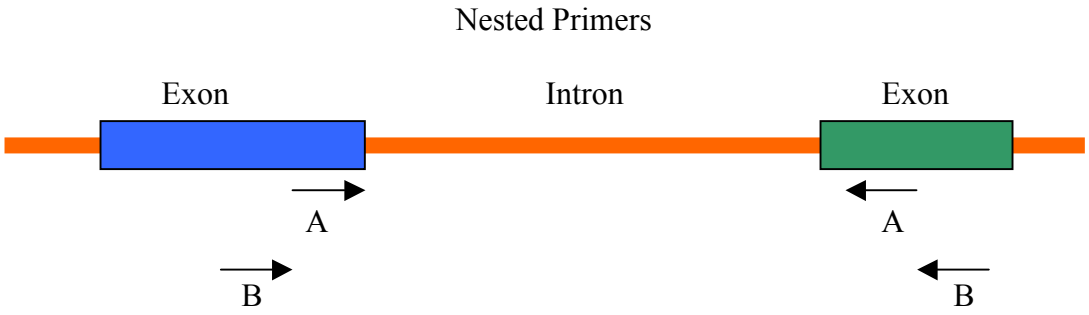


Figure 13: To promote the amplification of specific DNA corresponding to the intron, PCR of genomic DNA was performed with the primer set. The PCR was then diluted and used in a PCR with the primer set. Arrows "A" represent inner primers and arrows "B" represent outer primers. The goal of nested primers is to increase selectivity on during amplification in a PCR.

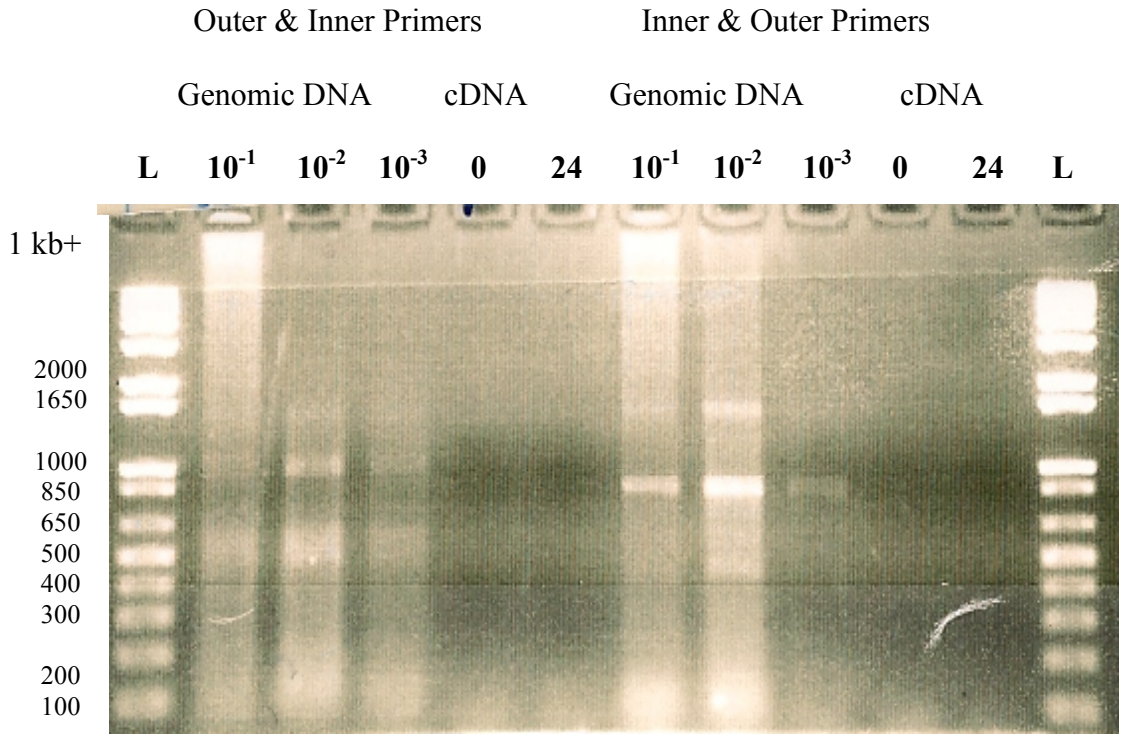


Figure 14: Different combinations of inner and outer primes were used to make sure the primers were good and to see if the amplification could be more specific and only amplify the heat shock protein. The genomic DNA, which includes an intron should be around 1500 bps and one of the bands is around the predicted size. "L" represents a 1 Kb+ ladder, and the numbers below each the genomic DNA represent the dilution of genomic DNA. The numbers below the cDNA represent cDNA that was extracted at 0 hours and 24 hours.

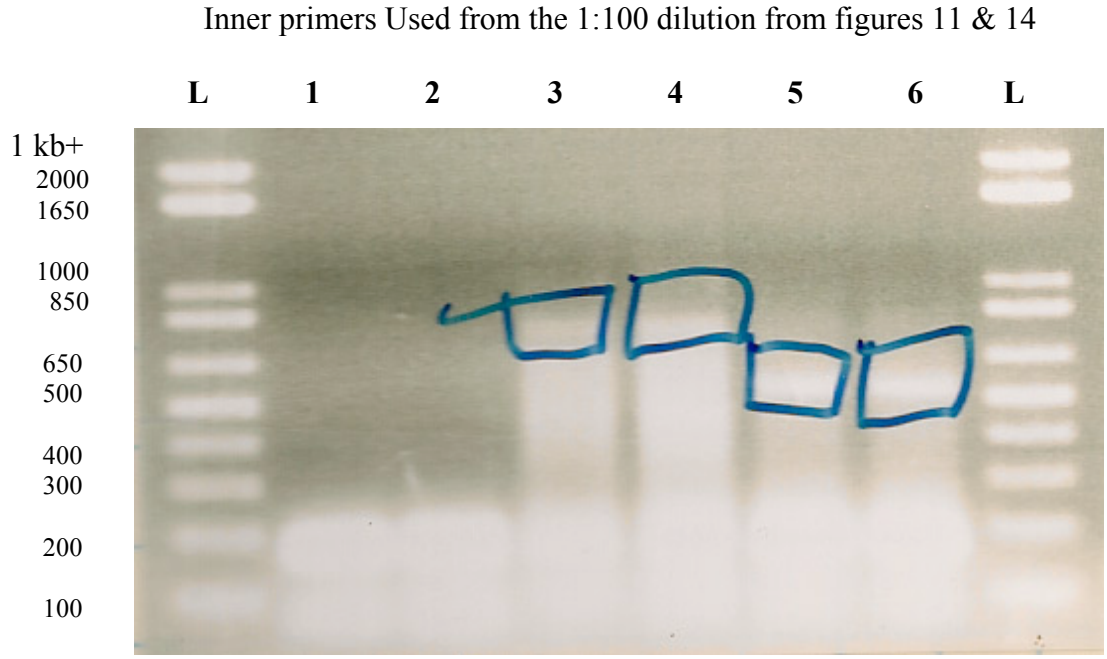


Figure 15: PCR products from figures 11 and 14 at 1:100 dilution were used and amplified again using just the inner primers. Lanes 1 and 2 represent the outer primers, lanes 3 and 4 represent the outer and inner primers, and lanes 5 and 6 represent the Inner and outer primers. The odd numbered lanes were at 1:50 dilution of the PCR product, while the even lanes were at 1:100 dilutions from the PCR product.

III.3 Discussion

Amplifying fragments from *Artemia* DNA was not originally a success based on the published sequence of the *Hsp26* gene published at the time. I expected to amplify a 1500 base pair fragment. However, a PCR of the genomic DNA using the outer primer only yielded a number of smaller bands (Figure 11). To determine what sequences were amplified in these reactions I extracted, cloned, and sequenced multiple bands from gels. I was unable to isolate the intron from the Hsp26 gene.

Some possible problems with this module were that fact that the genomic DNA was made from a different isolate than the one used to sequence the genomic *Hsp26* gene. Instead of a single 1500 base pair fragment, the genomic amplification produced a number of smaller fragments. There could have been a difference among sequences between species of *Artemia* from the NCBI data base and my samples obtained from the pet store. At the time, there was only one genomic sequence of *Hsp26* available to compare to. This would only really affect the outcome if there were polymorphism in the primer sequences that were chosen. In order to achieve greater success, one would need to create both a cDNA library and a genomic DNA library from the same pool of organisms. The primers were made to make sure they fulfilled the requirements, but the DNA sequence was very A-T rich, and designing primers with high specificity was difficult. The sequences obtained from NCBI could also have contained a few errors or even single-nucleotide polymorphism (SNPs), which could affect the ability of the primers to anneal and amplify.

More work needs to done to fine-tune this technique. Because of the limited amount of genomic sequences of *Artemia franciscana* available in the NCBI database at

the time, and only one genomic sequence of the heat shock protein, limited our control and provided uncertainty that we were using the same species. Trying this technique with other known sequences may have a great success rate because there may be more conserved regions for primer development. Conditions can be adjusted further for the PCR. Choosing primers that are in highly conserved coding regions may provide better results.

Students may be assessed along the way, making sure they are grasping important concepts and such as understanding the difference between genomic DNA and cDNA, properly developing primers based on a known sequence, comparing two different sequences looking for slight differences in both the coding and noncoding regions, the ability to make inferences based on their results, looking for answers to questions that may not be available, and hypothesizing their own explanations for their results. These are all important concepts that students should walk away with after completing this module. Test and quizzes can periodically be given throughout the module; depending on how much time the teacher spends on it. Students should also maintain a detailed lab journal that the teacher should collect and comment on as a form of formal assessment and then during the test have students answer multiple choice questions as well as essay questions with a defined rubric. Students with exceptional understanding should be able to give detailed explanations of the above-mentioned concepts, citing examples from their experiments and using accurate scientific terms in their responses. Students with satisfactory understanding may leave out many of the scientific terms and not fully site examples from their research. Students with an unsatisfactory understanding of the concepts will not be able to describe and explain all the concepts listed above and cite

little to no examples from their research.

Chapter IV: Teaching Module III – Measuring Gene Expression

IV.1 Introduction

Organisms do not express all their genes at once. This would waste a lot of energy and could potentially kill the organism. As an organism grows, different genes are expressed that facilitate development. Some genes are always expressed throughout an organism's life, however other genes may only be needed at specific stages of development. Besides development, there may be specific genes expressed in times of stress. These genes can code for proteins that may help an organism survive changes in their environment, such as change in temperature, oxygen content, and availability of food. A cDNA library only represent genes that were expressed under conditions that the organism was grown under at the time of harvesting. Trying to find genes that are only expressed at certain stages of the organism's life can help scientists better understand how the organism survives and develops.

Northern and Western blot analysis can be used to determine if the gene and protein are being expressed, but involves harsh chemicals and difficult techniques. An alternative to a Northern blot is to extract RNA at different time points and make a cDNA pool from each of those time points. This would allow scientists to compare the same gene to see if they expressed more, less, or not at all during these different time points. Extraction of RNA and converting it into cDNA is relatively difficult, expensive, and should be performed by the WSSP staff instead of the students. However, if the students are provided with the cDNA from the different time points they could perform a PCR to see how the gene they isolated is expressed at each time point. To do this, students need to first analyze their desired gene sequence and construct primers from both ends of their

clone to amplify the gene using a pool of cDNA. Students should be able to predict the size of the band and easily locate it on a gel. PCR amplification of the DNA fragment from the pool of cDNA will indicate if their gene was expressed. If the gene is not expressed, there would be no amplification. The intensity of the band may also inform students of the abundance of the gene during that specific period of the organism's life. If the students had access to different time points of the organism's life, they can use this technique to see their gene at different development stages.

IV.2 Methods and Results

Constructing a cDNA library is a little more challenging than creating a genomic DNA library. RNA is harder to work with, because it degrades much easier than DNA. This is because RNase is present everywhere and can easily break apart RNA.

To develop a protocol for this module, *Artemia* cysts were grown in a 1 liter flask on a shaker at low speed in *Artemia* hatch media (16). Aliquots of 10 ml were taken at 0, 4, 8, 24, and 48 hours from the flask, spun down, washed and placed into microfuge tubes and stored at -80 ^{0}C. A total dry mass of 0.5 grams of the organism was collected at each time point. Once 0.5 ug of *Artemia* were collected in a microcentrifuge tube, 400 µl of TES solution (10 mM Tris pH 7.5, 10 mM EDTA, 0.5% SDS) and 400 µl of acid phenol was added and vortex for ten seconds. The sample was incubated at 65 ^{0}C for 45 min, and vortex briefly 3 – 4 times during incubation. The sample was then placed on ice for five minutes, and then centrifuged for five min at 4 ^{0}C. The aqueous layer was transferred into a clean tube and 400 µl of phenol was added and vortex for 10 sec, then spun at top speed for 5 min at 4 ^{0}C. The aqueous solution was transferred into a new tube

and a DEPC treated solution of 40 µl of 3M NAOAc at pH 5 was added with 1 ml of ice cold 100% ethanol. This was placed at -80 ^{0}C for 30 min, and then spun for 5 min at 4 ^{0}C. The optical density was then measured and the pellet was dried and stored (Table 9).

To purify mRNA, the RNA was passed through an oligo-dT column that binds the poly-A tail of the mRNA while the other RNA's flow through the column. The mRNA was eluted from the column and an oligo dT primer was used with reverse transcriptase to convert the single stranded mRNA into cDNA using Clonetech® "Smart cDNA library construction kit"

Primers for tubulin and actin depolymerase were used in an attempt to amplify those genes from the cDNA at different time points using PCR (Table 10). Both of those genes are essential for viability and should be produced throughout its life cycle. A cDNA pool that was prepared by Marty Nemeroff to construct the *Artemia* cDNA library was used as a control to make sure my cDNA was prepared correctly (Figure 15). The intensity of the PCR bands indicates the level of expression. A more intense band represents greater expression of the gene than a time point with a less intense band. Unfortunately, the results varied too much and there was no consistency with the levels of expression between all the experiments. For example, separate PCR's show different intensities for the same time points (Figure 16). If these experiments were performed with a real-time PCR, results may be more consistent and reliable. Dilution of the DNA sample may also give better results with a regular PCR. High school students would not be able to perform real-time PCR at their schools.

Table 8: Optical Density of RNA Extraction

	Optical Density at 260 nm
Time of sample taken	
0 hours	2.2 µg/µl
4 hours	2.0 µg/µl
8 hours	3.2 µg/µl
24 hours	21.1 µg/µl
48 hours	1.152 µg/µl

Table 8: The O.D. of the 5 samples taken was performed at a wavelength of 260nm. The formula that was used for calculating the O.D. was = (O.D)(dilution)(40)/1000. The dilution was 300 fold and 40 represented the RNA standard. The sample taken at 24 hours may have not been completely purified or possible contaminated, because the OD was 7 – 10 times higher than the other samples. The 24 hours sample OD is extremely high and could either represent contamination or the RNA was not isolated correctly.

Table 9: List of Oligonucleoide Primers

Oligo Name	Gene Name	Accession #	Oligo sequence
1. Tubulin	beta-tubulin	EF660894	5' CAT GTT GGC CAA GCC GGT GTC C
			3' GTC AGA GTG CTC CAA GGT GGT ATG AG
2. Thiore P.	Thioredoxin Peroxidase	EF203465	5' CTG TTG GGT GGG GTG AAG AAG C
			3' CCA GCT GGG CAT ACT TCG CCG
3. Actin D	Actin Depolymerase		5' GGC TTC TGG AGT TCA AGT AGC TGA TG
			3' CGC GAA GTC TTT CTT CCA CGG C
4. Cyclo a	Cyclophilin a	EF203466	5' GCA CCC ACG AAA AGG GAT TTG GC
			3' CGG AAC TTT TCA ACG CGG TAA TGC C
5. Putative a	Putative Adenylate Cyclase	EF062480	5' GGA AAT GAT GGC CCA GAA TGC GG
			3' GCA CAA CCC CAC AGA TCT TAG TTC C

Table 9: This lists the 5 different genes that were selected from published sequences by students from WSSP on the NCBI website.

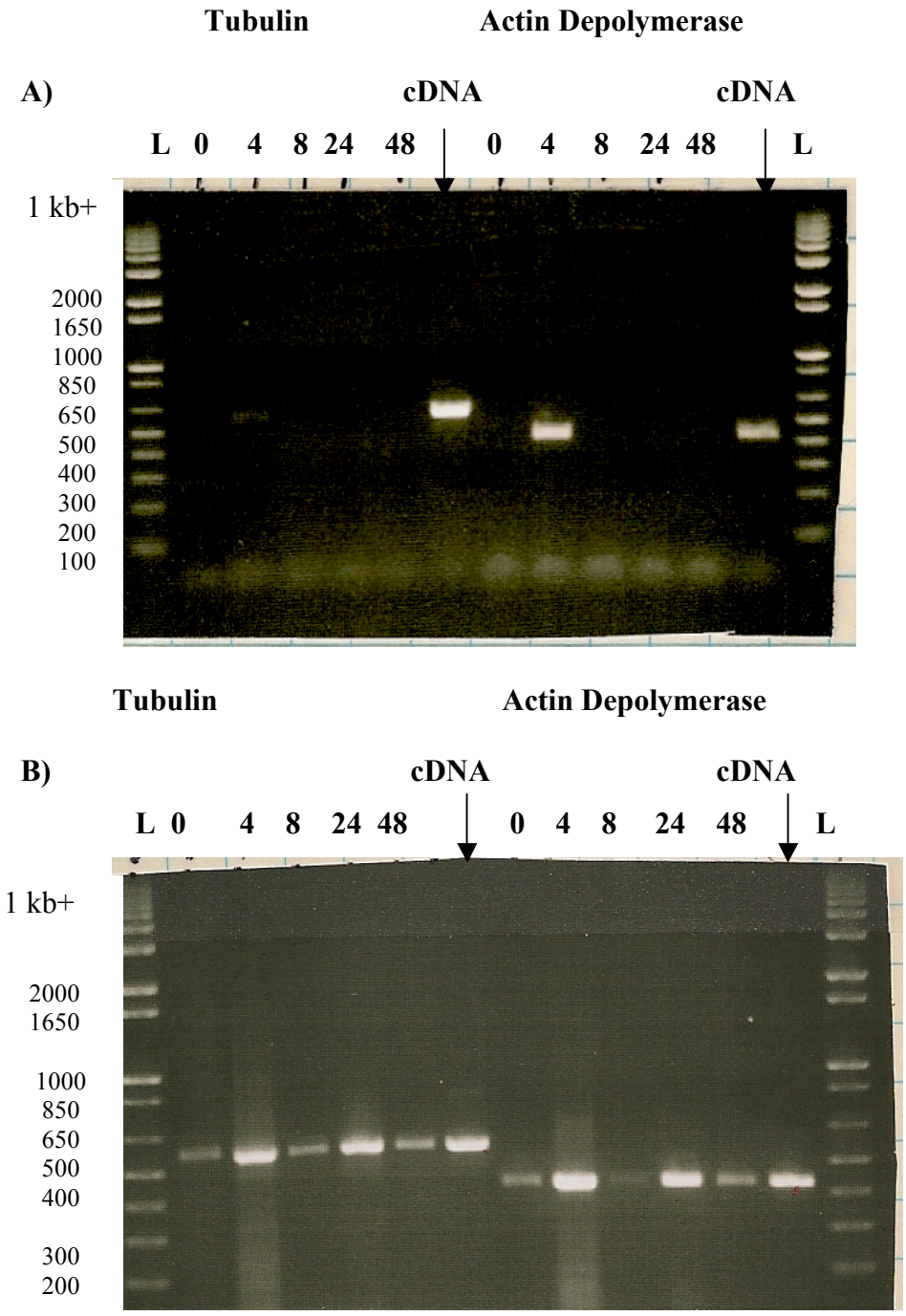


Figure 16: Tubulin and actin depolymerase were amplified from cDNA made at different time points, 0, 4, 8, 24, and 48 hours. "L" represents the 1Kb+ ladder used and the "cDNA" was cDNA made from adult *Artemia* by Marty Nemeroff as a control. A) Tubulin shows slight expression at 4 hours and in adult *Artemia*. Actin depolymerase shows stronger expression in 4 hours as well as in the adult cDNA. The predicted size of tubulin is 576 bases long and the actual size according to this is around 575 bases long. Actin depolymerase was predicted to be 421 bases long and the length is 450 bases long. B) Both genes are expressed, but hours 4 and 24 are expressed the most.

IV.3 Discussion

Measuring gene expression is a significant step in understanding an organism. One critical aspect of molecular biology is finding when a particular gene is expressed during the life of the organism. Certain genes are only expressed when the organism is under stress. The proteins coded by these genes may play a very important role in the organism, and the knowledge about the expression of these genes may have the potential to help cures diseases and better humanity.

The main problem with the protocols I tried to develop was the inconsistency and the difficulty in identifying the expression level by PCR. Trying to verify results was difficult, because there was so much variation when looking at the intensity of the bands. After 30 cycles in the thermocycler, the amplification process may have reached a plateau and expression levels normalized. Possible corrections could be to run fewer than 30 cycles or diluting the template cDNA further. More work and time could have been invested in fine-tuning the elongation and cycles for optimal PCR conditions, as well as normalizing the cDNA. However, since each gene and primer sets the students use will have different parameters, this may be too much work for a high school setting. However, given the time in an advanced research project this may be feasible.

A rubric for this section should focus on the reasons why genes are not always expressed and controls an organism uses to regulate gene expression. A proficient student should be able to explain that not all genes are expressed and some reasons why. An advanced proficient student should also add how genes are regulated and be able to site examples of gene regulation.

V.I Conclusion

The lab modules developed for students to perform experiments facilitates the opportunities and possibilities of several of the goals of science education by allowing students to perform investigations over an extended period of time and applying their results of experiments to scientific arguments and explanations (7). The experiments developed in the WSSP act as inquiry based learning, enabling students to learn by doing while using models based on genuine research programs. Students need a breadth of background information when conducting experiments in order to draw conclusions and make inferences (7). Technology also plays a large role in scientific research and discovery. Most research labs have access to the latest and best equipment, while high schools have limited funds and resources. The WSSP lends out the equipment needed to perform the experiments, enabling students to collect information and test hypothesizes. The WSSP also covers many of the life science content standards for the Nation Science Education Standards, such as understanding the cell, molecular basis of heredity, biological evolution, interdependence of organisms, and organization in living things.

Assessing student progress in a standard and conventional high school lab is fairly simple, because there is usually only one correct answer. In contrast, when performing an inquiry based experiment, assessing students becomes more challenging because there may be many acceptable outcomes. Following with the National Science Education Standards, the labs in the WSSP align with their goals to determine the student's initial understandings and abilities, to monitor student progress, and to collect information to grade student achievement. Achievement does not have to be measured in right or wrong answers, but students can reflect on what they have learned in a journals or lab

notebooks. Journals can be collected and graded with accordance to designed rubric. A rubric should be developed that represents standards for a certain assessment. Teachers should design the rubric based on their given population of students, what they would want the students to walk away with, and the overall goal of producing scientific literate adults (7). Rubrics should look for key terms, logical reasoning, understanding of concepts and linking concepts together and not just the regurgitating of facts. Making rubrics available to students before hand allows them to have an idea what is expected of them. Results should be classified as exemplary (entirely correct), satisfactory (partially literate), or inadequate (7). All goals should align with the schools curriculum and the state or national science standards.

Of the three modules, subcloning was the most successful, and has been performed several times by the WSSP students and teachers. Even though the educational value is very high for subcloning, it is challenging and students will have to put in an effort to be successful. Even if schools do not have access to all the equipment and reagents, a virtual subcloning experiment can be performed and students can still figure out what enzymes need to be used and the conditions for digestion.

The genomic cloning module also has potential. The WSSP recently switched organisms and libraries and may have better success with this module. This module can be further developed and possible used. The gene expression module proved to be the most challenging module to perform and may not be suitable for high schools until new and simpler techniques are developed for gene expression. Schools will need to modify their curriculum in order to fit any of these modules in. Teachers will have to move away from teaching by lectures and teach the same material using hands on inquiry-based

lessons like these modules.

Chapter VI: References

1. Griffith, Frederick (1928-01). **The Significance of Pneumococcal Types**. The Journal of Hygiene 27 (2): 113–159.

2. Avery, Oswald T.; Colin M. MacLeod, Maclyn McCarty (1944-02-01). **Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types: Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type III** . *Journal of Experimental Medicine* **79** (2): 137–158.

3. Vershon, A. **Chapters 1, 2, 6, 9 12, 13.** *WSSP-08 Lecture and Lab Files* 2008

4. Volume 7 Number 6 1979 Nucleic Acids Research (first miniprep)
A rapid alkaline extraction procedure for screening recombinant plasmid DNA
H.C.Birnboim+ and J.Doly Laboratoire de G6n&tique MoleCculaire, Institut de Recherche en Biologie Moleculaire, F-75221 Paris Cedex 05, France
Received 3 August 1979 Volunme 7 Number 6 1979 Nucleic Acids Research

5. Chinn, C.A. & Mohltra B.A. **Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks.** *Science Education, 2002,* **86**(2), 175-218.

6. Lemke, J. L. **The missing context in science education: Science.** Paper presented at a multi-disciplinary symposium entitled, In search of inquiry, at AERA, Atlanta, GA. Arlington VA: ERIC Documents Service (ED 363511), 1994.

7. National Research Council. **National science education standards: Observe, interact, change, learn.** Washington, DC: *National Academy Press.* 1996

8. Charney, J., Hmelo-Silver, C. E., Sofer, W., Neigeborn, L., Coletta, S., & Nemeroff, M. **Cognitive apprenticeship in science through immersion in laboratory practices.** *International Journal of Science Education* 2007, **29**, 195 – 213.

9. Oxford Dictionary of Biochemistry and Molecular Biology.

10. Bell, R. L., Blair, L. M., Crawford, B. A, & Lederman, N. G. **Just do it? Impact of a science apprenticeship program on high school students' understandings of the nature of science and scientific inquiry**. *Journal of Research in Science Teaching*, 2003, **40**(5), 487–509.

11. Lawless, J. G., & Rock, B. N. **Student scientist partnerships and data quality.** *Journal of Science Education and Technology* 1998, **7**(1), 5–13.

12. Handler, A.T. and K.M. Duncan. **Partnership research projects: scientists can help improve pre-college science education without compromising research goals.** *The Science Teacher* 2006, **73**: 36-39.

13. Qiao Zhou, Juliana Brown, Andrew Kanarek, Jayaraj Rajagopal & Douglas A. Melton. **In vivo reprogramming of adult pancreatic exocrine cells to b-cells.** *Nature* 2008, **455**: 627-633

14. Pierce, Benjamin A. **Genetics: A Conceptual Approach**. New York City: *W. H. Freeman and Company*, 2008.

15. Toon, J. **Study shows China as World Technology leader.** *Georgia Tech. Research News & Publications Office Georgia Institute of Technology.* 2008

16. Langdon CM, Bagshaw JC, MacRae TH. **Tubulin isoforms in the brine shrimp, Artemia: primary gene products and their posttranslational modification.** *Eur J Cell Biol.* 1990 **52**(1):17-26.

17. Beerer, K., and Bodzin, A. **How to develop inquiring minds. District implements inquiry-based science instruction.** *Journal of Staff Development,* 2004 **25**(4), 43-47.

18. President Barak Obama Speech. **Remarks by The President at The National Academy of Sciences Annual Meeting.** National Academy of Sciences Washington, D.C. April 27, 2009

19. WestEd on the Internet. 2010. http://www.wested.org/cs/we/print/docs/we/home.htm