Conducting a PCR Reaction & DNA Sequencing to Identify Genes in a Plasmid

By: Thomas Hoke

BIOL 250

Fall 2020

Abstract

Throughout this lab there have been two main targets. The first was to educate students on the proper methods for isolating, identifying, manipulating, and analyzing a gene from a plasmid. The second goal was to conduct those methods to gain an understanding of how expression of a GFP gene is regulated by the pGLO plasmid and environmental factors. To isolate the pGLO plasmid a PCR was used, and the product was identified through DNA sequencing via BLAST analysis. Once it had been confirmed that correct piece of DNA had been isolated, that plasmid was manipulated via mixing with the ampR gene and arabinose-C or glucose sugars on LB agar plates with *Escherichia coli* (*E. coli*) bacteria and analyzed for expression of the GFP gene. The findings of this study have been that the GFP gene is expressed when it is in the presence of arabinose-C sugar. This is because the ara-C sugar activates the GFP gene for transcription through the araBAD promoter which proceeds the GFP gene in the pGLO plasmid. Through these results the students have gained an understanding of what genes make up a pGLO plasmid and how those genes can be expressed or inhibited based on their surrounding environments.

Introduction

Large time and capital investments are needed to undertake a polymerase chain reaction (PCR) and DNA sequencing, but also practice of the required procedure is) needed to train individuals so that they can undertake experiments or tests where it is needed. Human health represents a motivation which has driven many of the advances in today's scientific realm and has also taken a central role in setting the goalposts for this project. The analysis of whole pathogen genomes by DNA sequencing has been tipped to lead to the next wave of scientific advances in medical microbiology as it has allowed for researchers to describe and track strains that had caused an outbreak (Tagini and Greub 2017). However, the real strength of PCR and sequencing when it

comes to health applications are real-time tests that allow researchers to identify threats before they take large affect. The identification of avian schistosomes in Wisconsin and Nebraskan lakes has been done via PCR and sequencing assays that target the 18S & 28S rRNA genes and have allowed for rapid identification of waterbodies where this threat is present (Jothikumar et al. 2015). In this research the central question has been if it is possible to isolate and identify a GFP gene from a pGLO plasmid? If the GFP gene could be isolated then it is expected that expression would be observed when the gene was introduced into a conducive growth environment.

Methods

This semester long lab project had a series of experiments which were undertaken by students, the materials required and processes undertaken for each experiment will be detailed in this section.

Polymerase Chain Reaction (PCR)

To conduct a polymerase chain reaction an ice bucket was obtained and DNA template, DNA primers, master mix, and water was placed in it. The DNA primers were F1 which had a sequence of 5' atcgcaactctctactgtttc 3' and R1 which had a sequence of 5' tctgatttaatctgtatcaggc 3'. First 20 µl of water was added to the tube followed by 10 µl of the DNA template. 20 µl of the DNA primer were then added, they had been mixed before the lab, and finally 50 µl of the master mix. The thermal profile for the PCR machine was 4 minutes at 94C for denaturization followed 30 seconds at 50C for annealing and lastly 7 minutes at 72C for extension. The PCR products where then cooled to 4C and stored overnight.

Gel Electrophoresis

To create the gel for electrophoresis an agarose gel powder and buffer was obtained. 100 ml of 0.25x TAE buffer was added to 1 gram of agarose gel powder in a beaker and then put into a microwave without mixing. It was microwaved for 45 seconds and removed from the microwave. The gel was allowed to cool to the touch, then poured it into the gel tray. The comb was inserted and let solidify. Once the gel was prepared the DNA solution, which was the product of their PCR reaction, was diluted with loading buffer. 7 µl of loading buffer was mixed with 3 µl of the DNA solution. To prepare the gel for electrophoresis it was placed on a platform with the wells on the black electrode (-) side and TAE buffer was added until it filled the wells and covered the top of the gel. The instructor loaded the DNA ladder and then students loaded their mixtures into a well. The lid was placed on the electrophoresis chamber and leads were connected to the power supply with red to red and black to black. The power supply was turned on and the gel was ran for 25 minutes at 300V. The gel was then moved to the visualizer and illuminated with UV light so that a picture could be taken.

PCR Amplicon Purification and Introduction to Sequencing

To purify the PCR product first a 1.5 mL tube that contained 500 µl of binding buffer was obtained and labeled. The sample was centrifuged for 60 seconds at 13,000 rpm. Next 200 µl of DNA wash buffer was added to the spin filter column and centrifuged for 60 seconds at 13,000 rpm. This wash process was completed in this exact fashion one more time. A 1.5 mL microcentrifuge tube was obtained and labeled next. The spin filter column was then transferred to this tube while make sure that the column's tip did not contact the flow-through. While being careful not to touch the white filter, 30 µl of elution buffer was added to the center of the filter. The column was then rested for 1 minute to allow the rehydration and suspension of the DNA.

The sample was then centrifuged one more time for 60 seconds at 13,000 rpm. After this the DNA had moved out of the column and settled at the bottom of the microcentrifuge tube.

A nanodrop was used to measure the concentration of the DNA as well as the A260/A280 values. To complete this 2 µl of the purified product was measured out and the nanodrop was "blanked" using loading buffer. The nanodrop was then run and the concentration of the DNA was recorded as well as the A260/280 values. The sequencing portion of this lab was outsourced to Eurofins Genomics, but the reactions were assembled in the lab. The primers used for sequencing, GLO F1 & R1, were the same ones used for the PCR reaction, with only one primer being used per sequencing reaction. To setup these reactions water was added to dilute the DNA stock to a level of 30ng/µl, after which 5 µl of the prepared solution was added to each tube along with 5 µl of one primer or the other.

BLAST Analysis of DNA Sequences

The results from the two sequencing reactions where downloaded online after being processed by the lap. There was one forward and one reverse sequence so the two needed to be stitched together for the later part of the lab. Each sequence was opened in SnapGene Viewer and copied to a text editor program where it was converted to FASTA format. Once both sequences had been saved as individual FASTA files they were combined using the site http://doua.prabi.fr/software/cap3 (accessed 9/30) which turned the two sequences into a Contig. The next step was to conduct BLAST analysis which was done using the site http://blast.ncbi.nlm.nih.gov/Blast (accessed 9/30). The nucleotide BLAST tool was used for this lab and the previously created Contig pasted into the box that was marked for it.. Once the search was completed a list of nucleotide sequences that matched closely with their search was

displayed with the “score” and “e-value” being the determining factors for how well the sequences matched. The top results of this search were captured for the results of this lab.

Genetic Engineering Using pGLO

To test for GFP expression two main samples were created, a -pGLO & +pGLO solution, which were then spread across agar plates to grow colonies. To do this first 0.25mL of 50mM $CaCl_2$ was pipetted into two microtubes, one labeled -pGLO and the other +pGLO, after which both were placed on ice. A sterile loop was then used to take a colony from a starter plate and add it to both tubes. A pGLO plasmid DNA solution was obtained and examined under UV light after which 10 ul was added to the +pGLO tube. Another 10 ul of the 50mM $Cacl_2$ solution was added to the -pGLO solution. Both tubes were then incubated for 10 minutes in ice while four agar plates were labeled based on if they would have the DNA control or not along with the presence of amp or ara genes. After incubation was complete both tubes were set in a water bath of 42C for 50 seconds and then returned to ice for 2 minutes. A pipet was used to add 250 ul of LB nutrient to each tube and then both were placed in a 37C water bath for 10 minutes and then 200 ul was pipetted onto each of the appropriate plates. A sterile loop was used to spread the liquid across the whole surface of agar. Lastly the plates were taped together and placed upside down in a 37C incubator for 24-48 hours.

Analysis of pGLO Sequences Using SnapGene Viewer

For the analysis of the pGLO sequence the pGLO plasmid file was downloaded off canvas. The SnapGene viewer program was opened by students and a new circular DNA file was created using the pGLO sequence. The “Detect Common Features” box was checked allowing the names of features to be written down along with their coordinates and expression direction. Once this information was gathered the ori(s), protein-coding genes, and promoters (7 features) were

selected. The next window showed the circular DNA and the “choose enzyme set” button was used to remove all enzymes. The primer sequences from week 2 were copied with forward primer being added to the 5’ end of the GFP and the reverse primer added to the 3’ end. The resulting map of the pGLO plasmid was screenshotted for the results portion of this lab.

Prokaryotic Gene Induction with pGLO

For the last test of GFP expression 4 LB-AMP plates, 1 vial of HB101, 1 vial of L-Arabinose, 1 vial of D-Glucose, and 1 vial of sterile water were obtained. The plates labeled as follows, according to what sugar would be added, pGLO, pGLO+ARA pGLO+GLU, and pGLO+ARA+GLU. Next, 4 sterile microcentrifuge tubes were labeled with the same names. Each tube was prepared with mixture totaling 250 ul of the sugars labeled. Each of the volumes from the tubes were micropipette up and down three times and then spread over the plate with sterile loops after being pipetted onto the plate. The plates were then placed in the incubator, set at 37°C, overnight.

Results

The confirmation that the polymerase chain reaction has been successful was done by gel electrophoresis (Figure 1) through which we determined the size of the amplicon was approximately 1000 base pairs (bp). The concentration of the amplicon was determined to be 150 ng/ul after being processed by a nanodrop. After adjusting the concentration to 30 ng/ul by mixing the product with H20 two tubes were prepared with a volume of 10 ul in each tube, 5 ul of primer and 5 ul of DNA, and sent to a lab for sequencing. The results of the sequencing were obtained via a Chromatogram (Figure 2) which allowed verification of the quality of the DNA strand. A Contig of the sequence (Figure 3) was assembled and BLAST analysis was then

conducted. It was determined that a pGLO gene had been isolated due to the correct open reading frame being present and the top match, 97% alignment, of the BLAST being a GFP cloning vector (Figure 4). To analyze if the pGLO plasmid was isolated and present in its full form a map (Figure 5) of the plasmid was created which showed the location of the oris, protein-coding genes, and promoters. After LB agar plates with amp and ara genes had been created and incubated to test for GFP expression the results were captured with a picture (Figure 6) showing that expression occurred on the plate where ampR and ara genes were present with pGLO. The number of colonies observed during this test as well as expression of GFP was recorded in (Table 2) as well. Another test for GFP expression was done with arabinose and glucose sugar being added to pGLO on LB agar plates and these results were captured via a table (Table 1) showing the number of colonies present and expression which only occurred on the pGLO + ARA plate.

Discussion

The central goal of this work was to isolate and identify a GFP gene which was successfully completed. The hypothesis of this work was supported as well since the GFP was expressed when introduced into a growth environment which arabinose sugar and ampR gene. Evidence of this is displayed in (Figure 6) with plate #1, which contained the above mentioned sugars and genes, displaying a visible glow when all other plates did not. Due to the students conducting this experiment being largely inexperienced in these procedures it took many hours to isolate, amplify, and analyze the correct piece of DNA. In the work by Jothikumar et al. (2015) the researchers had been able to conduct real-time tests which provide results that are applicable to real-world time sensitive scenarios. Another study by (Reuter et al. 2013) used rapid sequencing

to trace *E. coli* and other bacterial outbreaks with rapid testing, showing more real world applications and proof that given the correct resources and procedures rapid isolation and testing can be completed. The next step for this research would be to speed up the procedures which have been practiced thus far and determining if any time/cost saving measures can be implemented.

Figures/Tables

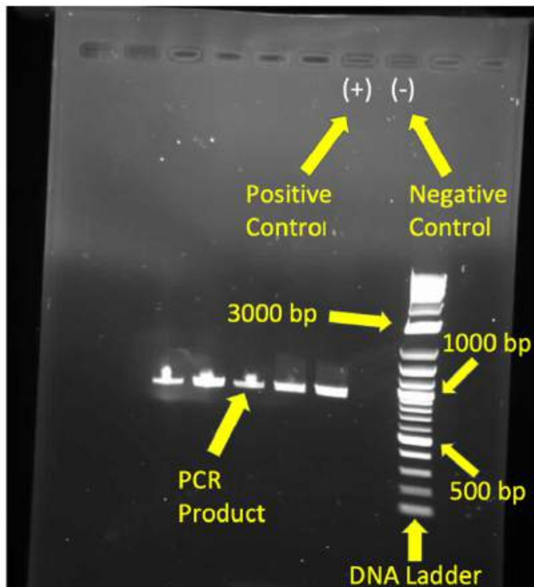

Figure 1. Image From Gel Electrophoresis of PCR amplicon

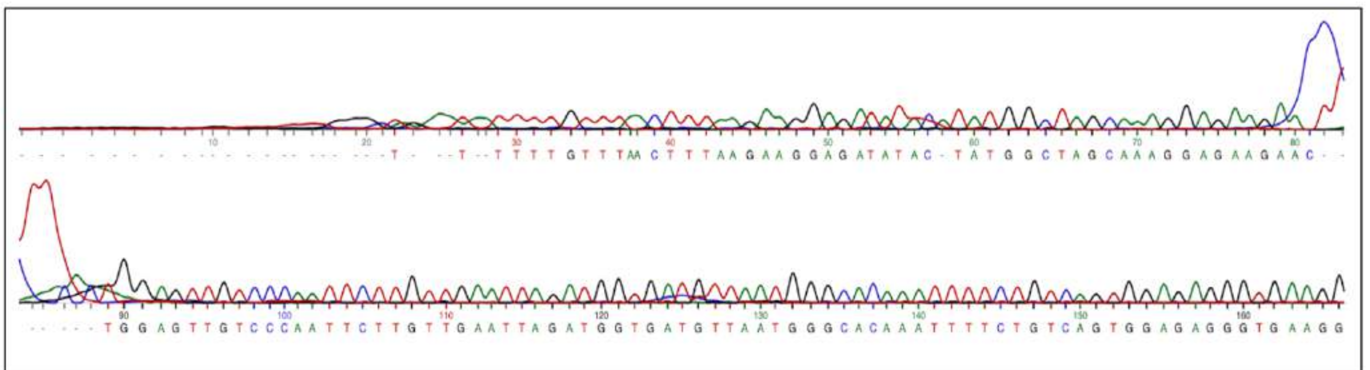

Figure 2. Chromatogram of a strand of pGLO DNA = Adenine = Cytosine = Guanine = Thymine

```
NNNNNNTNNACCNAACCNGGTNNCCCCCGNTNNTAAAAGCATTNNNNNACNAAGNGGGNN
CNANNCCATGACAAAAACGCGNANNAAAAAGNNNCTANATCNCGGCNGAAAAGTCCACAT
TGATTATTTGCNNGGCGTCACACTTTGCTATGCCATAGCATTTTTATCCATAAGATNAGC
GGATCCTANNGACGCTTTTATCGCAACTCTCTACTGTTTCTCCATACCCGTTTTTTTGGG
CTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGCTAGCAAAGGAGAAG
AACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACA
AATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCTACATACGGAAAGCTTACCCTTAAAT
TTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCTCTT
ATGGTGTTCAATGCTTTTCCCGTTATCCGGATCATATGAAACGGCATGACTTTTTCAAGA
GTGCCATGCCCGAAGGTTATGTACAGGAACGCACTATATCTTTCAAAGATGACGGGAACT
ACAAGACGCGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATCGTATCGAGTTAA
AAGGTATTGATTTTAAAGAAGATGGAAACATTCTCGGACACAAACTCGAGTACAACTATA
ACTCACACAATGTATACATCACGGCAGACAAACAAAAGAATGGAATCAAAGCTAACTTCA
AAATTCGCCACAACATTGAAGATGGATCCGTTCAACTAGCAGACCATTATCAACAAAATA
CTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCGACACAATCTG
CCCTTTCGAAAGATCCCAACGAAAAGCGTGACCACATGGTCCTTCTTGAGTTTGTAACTG
CTGCTGGGATTACACATGGCATGGATGAGCTCTACAAATAATGAATTCGAGCTCGGTACC
CGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCTGTTTTGGCGGATGAGAG
AAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAGCGGTCTGATAAAACAGAAT
TTGCCTGGCGGCAGTANCGCGGNGGTCCCACCNGACCCCNATGCCGAANCTCANAAANTG
AAACNCCGTANCNNCGATGGTAGNGNNGGGGNCNTCCCCCATGCGAGANTAGNNNNGNCN
NGNNATCAAATAAAANNNAANNNNNNNNNCNAANGANTGNNNNNNNTTTTNNTCNNNNNN
TTNGNCGGNGNANN
```

Figure 3. DNA sequence contig = F1 Primer Sequence

= GFP Gene Open Reading Frame = R1 Primer Sequence

Cloning vector pBAD-GFPuv, complete sequence

Sequence ID: U62637.1 Length: 5371 Number of Matches: 1

Range 1: 1124 to 2305 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
2026 bits(1097)	0.0	1153/1190(97%)	11/1190(0%)	Plus/Plus

```
Query  66    CCATGACAAAAACGCGNANNAAAAAGNNNCTA-NATCNCGGCNGAAAAGTCCACATTGAT  124
Sbjct  1124  CCATGACAAAAACGCGTA-ACAAAAGTGTCTATAATCACGGCAGAAAAGTCCACATTGAT  1182
Query  125   TATTTGCNNGGCGTCACACTTTGCTATGCCATAGCATTTTTATCCATAAGATNAGCGGAT  184
Sbjct  1183  TATTTGCACGGCGTCACACTTTGCTATGCCATAGCATTTTTATCCATAAGATTAGCGGAT  1242
Query  185   CCTA-NNGACGC-TTTTATCGCAACTCTCTACTGTTTCTCCATACCCGtttttttGGGCT  242
Sbjct  1243  CCTACCTGACGCTTTTTATCGCAACTCTCTACTGTTTCTCCATACCCGTTTTTTTGGGCT  1302
Query  243   AGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGCTAGCAAAGGAGAAGAA  302
Sbjct  1303  AGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGCTAGCAAAGGAGAAGAA  1362
```

Figure 4. BLAST analysis top match alignment

Score/Identities =

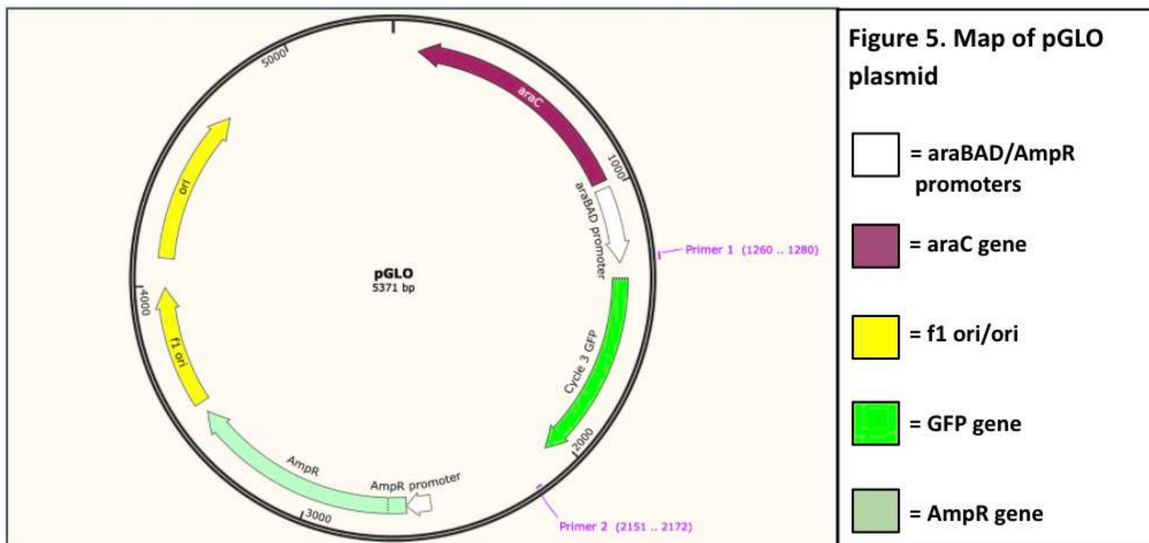

Figure 5. Map of pGLO plasmid

= araBAD/AmpR promoters

= araC gene

= f1 ori/ori

= GFP gene

= AmpR gene

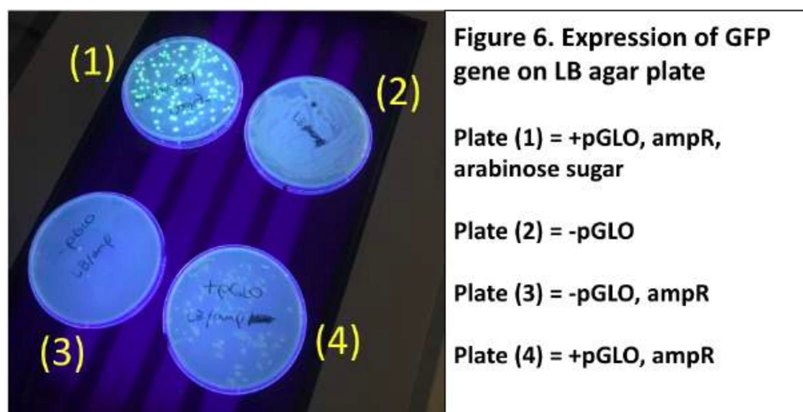

Figure 6. Expression of GFP gene on LB agar plate

Plate (1) = +pGLO, ampR, arabinose sugar

Plate (2) = -pGLO

Plate (3) = -pGLO, ampR

Plate (4) = +pGLO, ampR

Table 1. Colonies present and GFP gene expression in differing growth mediums

	pGLO	pGLO + ARA	pGLO + GLU	pGLO + ARA + GLU
Number of Colonies Observed	12	27	21	11
Expression of GFP	No	Yes	No	No

Table 2. GFP expression in transformed bacteria on LB agar plates

	+pGLO + ara + ampR	+pGLO + ampR	-pGLO	-pGLO + ampR
Number of Colonies Observed	47	31	0	0
Expression of GFP	Yes	No	No	No

Literature Cited

Jothikumar, N., Mull, B. J., Brant, S. V., Loker, E. S., Collinson, J., Secor, W. E., & Hill, V. R. 2015. Real-time PCR and sequencing assays for rapid detection and identification of avian schistosomes in environmental samples. *Applied and environmental microbiology*, *81*(12): 4207-4215.

Reuter, S., Ellington, M. J., Cartwright, E. J., Köser, C. U., Török, M. E., Gouliouris, T., ... & Parkhill, J. 2013. Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology. *JAMA internal medicine*, *173*(15): 1397-1404.

Tagini, F., & Greub, G. 2017. Bacterial genome sequencing in clinical microbiology: a pathogen-oriented review. *European journal of clinical microbiology & infectious diseases*, *36*(11): 2007-2020