The Sequencing of Arhodomomas Species Seminole Using PCR in Order to Better Analyze the Predicted Sequence Gap

GRP#9
Section#001

Dr. Patricia Canaan, Sierra Isreal, Sophia Short, and Elizabeth Shute

ABSTRACT

The objective of this project was to fill in a gap of a genomic sequence of Arhodomonas sp. Seminole. This is a bacterium that lives in salty, crude oil impacted soil in Seminole Co, Oklahoma. A program called Blastx was used to design a PCR primer which would allow the desired sequence to be made. Polymerase Chain Reaction or PCR, which is a biochemical technique used to amplify a specific piece of DNA by generating millions of copies of a particular DNA sequence done by splitting open cells to purify DNA and then concentrating the precipitation. After this gel electrophoresis was used to analyze the results. Our results showed that we did not choose the specific region in the DNA sequence that was needed.

INTRODUCTION

Arhodomonas Sp. Seminole was discovered at a crude oil site it Seminole County, Oklahoma. Arhodomonas sp. Seminole requires salt (NaCl) and grows on a wide variety of carbon sources (Canaan, 2014). Arhodomonas sp. Seminole can metabolize benzene, toluene, phenol, 4-hydroxybenzoic acid (4-HBA), protocatechuic acid (PCA), 12 and phenylacetic acid (PAA) as the sole sources of carbon at high salinity. To see how the bacteria functions in soil of high salinity the genome of the bacteria needs to have its DNA extracted from the bacteria and sequenced. After the extraction, the sequence was meant to come out complete but it did not. There was 750 pieces of genes in the Arhodomonas sp. Seminole genome, and not a single piece. With that many breaks, about 8-10% of the genes are broken between two pieces. The broken pieces of the genes are called contigs. We hypothesize that the end of Contig 590 is the head end that matches to the tail end of the Contig 510. The objective of this experiment is to be able to “fill in the gaps” (Canaan, 2014). The use of full fused contigs, a related protein, the PCR procedure, and an analysis of a Gel Electrophoresis will hopefully help to fill in the missing pieces.

MATERIALS AND METHODS

The objective of our experiment was to sequence the genomic DNA of the Arhodomonas sp. Seminole bacteria in order to fill the gaps of the DNA in order to better understand and predict the enzymes, traits, phenotypes, etc. To start this process our group was given two full length contigs, halo contig 590 and halo contig 510, as well as a related protein sequence relating to haloacid dehalogenase. With these contigs we used the Blastx program to try and align each contig with the related protein sequence. The halo contig 590 was found to be a head and the halo contig 510 was found to be a tail. Knowing this information we were given a fused contig sequence of the 510 tail and the 590 head. With this sequence we once again used BLastx in order to see if the fused contig was similar to the protein. From this we had to design our PCR primers by pulling out a sub sequence from the full fused contig sequence. This sequence had roughly 150-200 bases on each side of the gap. We ran this sub sequence that we created through Primerquest in order to find the most ideal forward and reverse primer sequences. Our sequences were the forward primer sequence of GCCATGACGGTGACACTGA and the reverse primer sequence of GTGGGTGAGCTGCTCGT. Our next step was to amplify these desired segments of DNA using a Polymerase Chain Reaction (PCR) in order to get millions of copies of our DNA segment in order to get a better analysis of the predicted sequence gap. In order for the PCR to work, we had to construct a Taq DNA Polymerase. To do this we used a micropipeter and disposable tips to add the reagents needed to construct the Taq Polymerase into a small test tube with 70 microliters of distilled water in it. First we carefully added 10 microliters each of the Taq buffer and the DNA nucleotide building blocks using the micropipeter. Next, five microliters of the Arhodomonas sp. Seminole gDNA was added. Lastly, two microliters each of our self-designed forward and reverse primers was added to the test tube along with 1 microliter of Taq Polymerase. Once our mixture was completed we had to insert a small amount into a gel cell. The test tube was inserted into a thermocycler to synthesize the DNA copies. Pictures of each gel cell were taken in order to see which groups obtained a PCR product, our group not having one.

RESULTS

The results from the PCR did not produce a visible product as shown in the image below. However, a search of the databases on NCBI for a protein that is similar to the protein that was in the genome produced some possible results. The results from the blast produced a protein that was highly similar to the protein in the strand of Arhodomonas. It was a hypothetical protein from
Arhodomonas sp. Aquaeolei. Hpothetical proteins are not proven to exist yet because there is not enough evidence to prove so. They are predicted proteins in a organism and the functions of them vary and are not able to be proven at this time (Lubec, et. all, 2014).The next highest proteins were all dehologenase haloacid which is an enzyme that belongs to the family of hydrolases, specific those acting on halide bonds in carbon-halide compounds (Zhang, 2003).

Figure 1

Figure 1: The image of the gel electrophoreses depicts the results of our PCR product and the red oval denotes where our protein should have appeared.

Figure 2

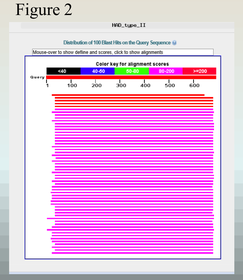

Figure 2: The most prominent enzyme was dehalogenase haloacid which is an enzyme that belongs to the family of hydrolases, specifically those acting on halide bonds in carbon-halide compounds.

DISCUSSION

The picture from the agarose gel electrophoresis, did not protein that was visible for us to be able to compare it to the ladder. We should have seen bands in our gel lane that indicated that we did, in fact, have a PCR product that might span the distance of the gap, but instead we saw an empty lane. We suspect that there are a few reasons why our proteins were not visible in our electrophoresis. The first speculation is that it might have failed because the sequence of nucleotide base pairs that was chosen for the primer was not the appropriate sequence. A possible solution for this problem would be to shift the position of the sequence that we made our primer from to see if this would be successful. If shifting the sequence of the primer did not work, another possibility would be to look for another region of the DNA sequence to choose a primer. Another possible explanation for the PCR not giving a result would be experimental technique. Since the materials that were needed for this experiment were transferred from test tube to test tube in microliters it is possible that some of the material was lost. Since the overall volume of the primer was so small, any amount lost could provide an ineffective result. Another technical error that could have been made would be how the gel was loaded. The pipette tip might not have been placed all the way into the gel well and all of the primer might have been lost which would have resulted in a blank gel.

The protein that came from the contigs did produce a related protein when we searched the NCBI databases. The protein that was most similar was the hypothetical protein from Arhodomonas sp. Aquaeolei. If hypothetical proteins could be proven to exist and their functions discovered then it could lead into many more explanations for why this bacterium is able to live in high amounts of salt and degrade carbons. It is a step to finding a way in cleaning up the crude oil sights that are hurting the environment.

There is still the possibility of finding more evidence from the protein that we had. Our project could continue in few future directions. The first path that we could take would be to repeat the experiment. The reason for repeating the experiment is to rule out the possible speculation of technical error. In order to do so, the loading of the gel should be watched closely to make sure it is done correctly and the loading of the samples into the PCR tube would need to redone to insure accuracy. In repeating the experiment and ruling out the possible technical errors we would know with certainty if our PCR reduced any solid evidence on the missing DNA in the gap. Another possible solution to receive a result from the PCR procedure would be to create new primer sequences. It could be shifted a little further from the gap in both upstream and downstream directions. This would look at more of the DNA and give a higher chance of allowing the PCR to span the gap in the DNA and produce results.

REFERENCES

• Canaan, Patricia. The Experiment. Presentation during Freshmen Research Biochemistry class, introduction to the experiment. August 26, 2014.
• Lubec, Gert, Leila Afjehi-Sadat, Jae-Won Yang, and Julius Paul Pradeep John. 2005. Searching for hypothetical proteins: Theory and practice based upon original data and literature. Progress in neurobiology 77, (1): 90-127.
•Zhang, Guofeng. 2003. "Diversification of Structure and Catalysis within the Haloacid Dehalogenase Enzyme Superfamily." Order No. 3093072, The University of New Mexico. http://search.proquest.com/docview/305312290?accountid=4117.