The Search For the Hidden Genome Sequence: Seminole Style

Group #15
Section #002

By: Cody Brooks, Mikayla Ott, Karoline Radka, and John Rawson
Instructor: Patricia Canaan

ABSTRACT

Arhodomonas sp. Seminole is a bacterium found thriving in crude-oil-impacted soil in Seminole County, Oklahoma. It requires salt to live and grows on a wide variety of carbon sources. After a soil sample was taken and analyzed scientists working on this project discovered that the genome sequence was incomplete resulting in 750 contigs. In an attempt to sequence these gaps we used PCR to replicate the space between two contigs. Upon receiving a positive result we used that information to align the sequences and discovered 1-Aminocyclopropane-1-carboxylic acid (ACC).

INTRODUCTION

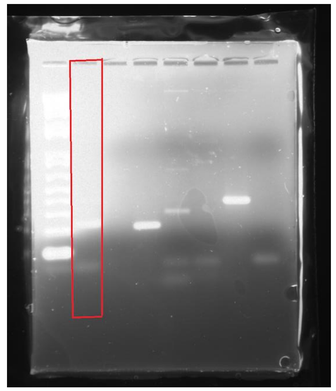
Arhodomonas sp. strain Seminole was isolated from a crude oil-impacted brine soil and shown to degrade benzene, toluene, phenol, 4-hydroxybenzoic acid (4-HBA), protocatechuic acid (PCA), and phenylacetic acid (PAA) as the sole sources of carbon at high salinity. Arhodomonas sp. strain Seminole was found in Seminole, Ok and requires salt to live and grows on a wide variety of carbon sources. A soil sample was taken from the area and was enriched to obtain a pure culture of salt-loving, aromatic hydrocarbon-metabolizing bacteria. We want to sequence the genome of the species. To do this, the cell was ruptured using glass beads and a vortex. This resulted in its single chromosome being broken up into several pieces. The chromosome was reconstructed using overlaps. However Arhodomonas sp. strain Seminole's genome was still broken up into 750 contigs, therefore needed to recreate the gap between two of these contigs.

MATERIALS AND METHODS

After we received our contigs, we used blastx to align them with the related protein also provided to us. Both contigs significantly matched the protein, so we used PCR to replicate the target DNA segment. We used PrimerQuest at IDT to determine the best forward and reverse primers. The forward primer was CACTGCCAGACCCATCAC and the reverse primer was CCTTCGGTGACCACAAGTC. We then created the solution to use for PCR. This was done by combining 70µL of dH2O, 10µL of Taq buffer, 10µL of nucleotides, 10µL of genomic DNA, 2µL of the forward primer, 2µL of the reverse primer, and 1µL of Taq Polymerase. This solution was placed in the thermocycler, which ran with the program which consists of an initial denaturation at 94° C for 30 sec, 30 cycles of denaturation at 94° C for 30 sec, annealing at 55° C for 30 sec, extension at 72° C for 1 min another 10 min at 72° C for the final extension. We mixed 8µL of this product with 2µL of tracking dye with a micropipeter, then inserted it into a well of 1% agarose gel for agarose gel electrophoresis. The PCR product was then sequenced, and we used ClustalW2 to align the products of the forward and reverse primers with the artificially fused contigs.

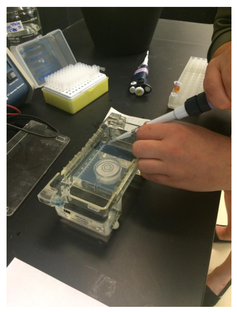
RESULTS

Our gel electrophoresis process resulted in a clear DNA product. The second well from the left illustrates our results. This shows that our PCR methods were done correctly. The picture can also be used to show the size of our DNA product, since we know that the further the band is away from the starting well, the smaller the product is.



When the products from our forward and reverse primers and our fused contigs ran through the clustal website, these are the results that we received. These tell us that our PCR products align with our fused contigs and that our two separate contigs are adjacent to one another.

>15F
TGGCCCGGTTTAATCCACCGACAGCCGTCTACGACCTGGGTCCGGCAACATTCTCTGAGTCGCGTCAT
GGGCGCGAACGTGGAGCTGGTGGACGAGGGTTTCGACATCGGCATCCGCGAAAGCTGGCGGCGCGC
CATCGAGGACGTGAAGGCGCAAGGTGGA
>15R_rc
CCGGCAGGTCGCGGCCGTGGCCGCGAAGCTCGGTCTAAGCTGCCGGCTCGTCCGGCAACATTCTCAT
GAGTCGCGTCATGGGCGCGAACGTGGAGCTGGTGGACGAGGGTTCGACATCTGCTCCGCGAAAGAA
GTATTTCC



DISCUSSION

Polymerase Chain Reaction (PCR) is the biomedical technology in molecular biology used to amplify a single copy or several copies of a piece of DNA across several orders of magnitude generating thousands to millions of copies of a particular DNA sequence. In order to obtain the proof that we had a product we used the agarose gel electrophoresis, a method used in biochemistry and molecular biology. The biomolecules are separated by applying an electric field to move the charged molecules through an agarose matrix and the biomolecules are separated by size in the agarose gel matrix. BlastX is a database that provides a variety of genome databases for different organisms. It links a protein sequence to a nucleotide sequence and translates them to genomic information. The program ClustalW2 is used to align multiple sequences for DNA or proteins. It helps find the best match for certain sequences and lines them up so the similarities and differences can be seen. 1-Aminocyclopropane-1-carboxylic acid (ACC) a is disubstituted cyclic alpha-amino acid in which a three membered cylclopropane ring is fused to the C(alpha)-atom of the amino acid.

REFERENCES

Dalvia, S., Nicholson, C., Njar, F., Roe, B. A., Canaan, P., Hartson, S. D., Fathepure, B. Z. (2014, August 15) Arhodomonas sp. Strain Seminole and its Genetic Potential To Degrade Aromatic Compounds under High-Salinity Conditions. *American Society for Microbiology.* Retrieved from http://aem.asm.org/content/80/21/6664.short