Annotation of the Pseudomonas Aeruginosa Genome at Locus Tags T223_00110, T223_00115 and T223_00125

Paige Phillips, Emily Leonard, Elizabeth LaFever, Natalie Masten, Gabrielle Kashorek, Alicia Faville, and Mary Keymel
Avon Central School and The Western New York Genetics in Research and Health Care Partnership

SEPA SCIENCE EDUCATION PARTNERSHIP AWARD
Supported by the National Institutes of Health

University at Buffalo

Abstract

A group of 3 genes from the microorganism *Pseudomonas Aeruginosa* (T223_00115. T223_0010, and T223_0025 were annotated using the GENI-ACT website. The bacteria, *Pseudomonas Aeruginosa* was closely studied in a series of tests that broke down and revealed the complex amino acid sequences and functions of the genes. Our research gathered information on DNA coordinates, protein sequences, Blast, CDD, T-Coffee, Gram Stain, TMHMM, SignalP, PSORT, Phobius, IMG Sequence, TIGRfam, Pfam, and PDB. The gene annotation showed that loci T223_0010,T223_0015, T223_0025 are mostly cytoplasmic proteins. research concludes that all the genes were correctly annotated by the database.

Introduction

During our time with GENI-ACT we have been studying our bacteria called *Pseudomonas Aeruginosa. Pseudomonas Aeruginosa* is a Gram-negative rod shaped bacteria. The bacteria can survive in a wide variety of conditions with minimal nutrients. *Pseudomonas* can be found in soil, water and a widespread variety of vegetation. According to Friedrich (2016), *Pseudomonas aeruginosa* has become an important cause of gram-negative infections, especially in patients with weakened immune systems. This bacteria is associated with ear infections, skin rashes, eye infections, infections in blood, and pneumonia (Friedrich,2016. *Pseudomonas Aeruginosa* is very life threatening.

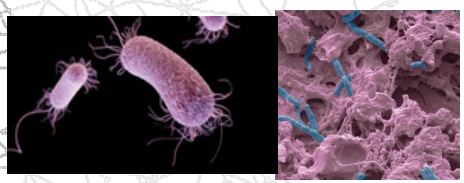

Figure 1. A three dimensional computer generated image of *Pseudomonas Aeruginosa*. Image: James Archer, Centers for Disease Control

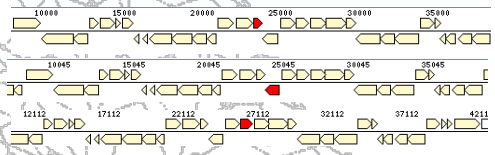

Figure 2. The relative positions of the genes for locus tags T223_0010,T223_0015, and T223_0025

Methods

Modules of the GENI-ACT (http://www.geni-act.org/) were used to complete the *Pseudomonas Aeruginosa* genome annotation . The modules are described below:

Modules	Activities	Questions Investigated
Module 1- Basic Information Module	DNA Coordinates and Sequence, Protein Sequence	What is the sequence of my gene and protein? Where is it located in the genome?
Module 2- Sequence-Based Similarity Data	Blast, CDD, T-Coffee, WebLogo	Is my sequence similar to other sequences in Genbank?
Module 3- Cellular Localization Data	Gram Stain, TMHMM, SignalP, PSORT, Phobius	Is my protein in the cytoplasm, secreted or embedded in the membrane?
Module 4- Alternative Open Reading Frame	IMG Sequence Viewer For Alternate ORF Search	Has the amino acid sequence of my protein been called correctly by the computer?
Module 5- Structure-Based Evidence	TIGRfam, Pfam, PDB	Are there functional domains in my protein?

Results

T223_00115

The GENI-ACT top hit was Quinone oxidoreductase. This protein has no helixes predicted and has no cleavage site. There is also no signal peptide present. According to PSORT B, the predicted location of my protein is in the cytoplasm of a cell based on a score of 9.97. The E-value is 1.78e-113. My protein is associated with esophageal cancer. My protein is found widely in plants, in the general (benzoyl). Commercially, quinones are used in making dyes, tanning hides and in photography. It is also in many types of tumors, including the lung, ovary, adrenal gland, thyroid, liver, colon, breast, and pancreas. My protein is very well conserved through my whole sequence logo. The DNA coordinates of my protein are 24555..25532.

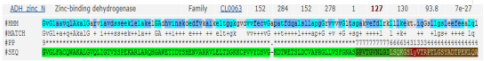

Figure 3. This is the Pairwise Alignment of my protein. The pairwise alignment shows matches between the model and your protein sequence

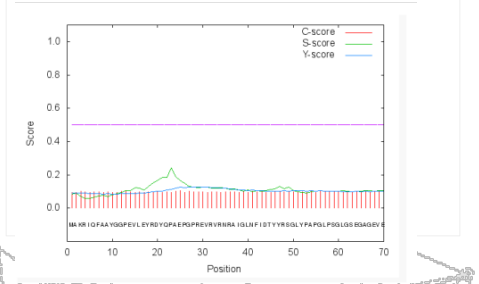

Figure 4. This graph is a signal peptide graph. The graph is showing that there is not a signal peptide in my protein.

T223_00125

The GENI-ACT gene name top hit was Shikimate 5-Dehydrogenase. Not only is this gene found in the bacteria *Pseudomonas aeruginosa* but it's also found *in Acinetobacter baumannii.* This gene's E-value is 4.65e-74. There are no predicted helices present, no signal peptide present, The protein is most likely found in the cytoplasm.

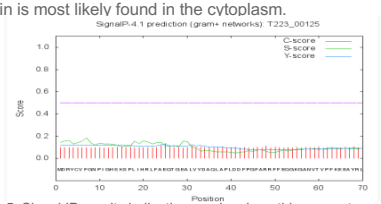

Figure 5. Signal IP results indicating no signal peptide present

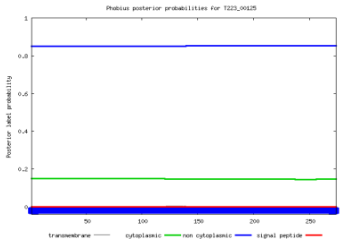

Figure 6. Phobius results showing lack of signal peptide or transmembrane helixes.

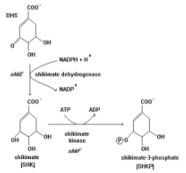

Figure 7. This is the function of our protein shikimate 5-dehydrogenase. The pathway links the metabolism of carbohydrates to the biosynthesis of aromatic compounds and is essential for the biosynthesis of aromatic amino acids and other aromatic compounds in bacteria, eukaryotic microorganisms and plants. aromatic: chemistry of an organic compound

T223_00110

The GENI-ACT top hit was for the a tRNA threonylcarbamoyladenosine biosynthesis protein. This genes E-value is 3e-81. This protein has no helices. Also, there is no signal peptide. Therefore this protein is most likely going to be found in the cytoplasm of a cell. This protein was also located in the organism *Enterobacter cloacae.* This organism is a bacteria that causes infections. These infections include skin infections, lower respiratory tract infections, and many more infections.

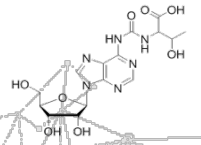

Figure 8. N6-threonylcarbamoyl is a nucleoside found in transfer RNA

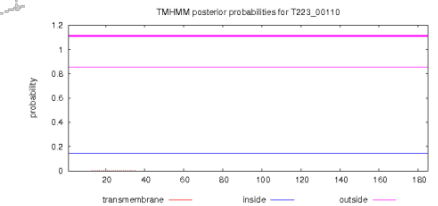

Figure 9.THMM results showing no transmembrane helices

Telomere Recombination:

My research showed that this protein is involved in Telomere regulation. Telomere Recombination occurs on the ends of chromosomes. Everyone (humans and animals) is born with telomeres that shorten naturally over time. Telomeres in humans can shorten at faster rates due to smoking, extreme physical activity, unhealthy diets and stress.

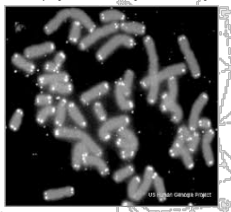

Figure 10. Human chromosomes capped with telomere recombination.

Conclusion

The GENI-ACT gene products did not differ from the proposed gene annotations for any of the gene loci investigated. The genes appear to be correctly annotated

Gene Locus	Geni-act Gene Products	Proposed Annotation	Protein Function
T223_00110	N6-threonylcarbamoyl adenosine	N6-threonylcarbamoyl adenosine	Found in tRNA, responsible for all ANN codons of life.
T223_00115	Quinone oxidoreductase	Quinone oxidoreductase	Involved in the metabolism and is related to tumors in organs.
T223_00125	Shikimate 5-dehydrogenase	Shikimate 5-dehydrogenase	Links the metabolism of carbohydrates to the biosynthesis of aromatic compounds.

References

Friedrich, Marcus ,"Pseudomonas aeruginosa Infections" Medscape 6 December 2016 Web. March 2017
https://www.cdc.gov/hai/organisms/pseudomonas.html

Acknowledgments

Dr. Rama Dey-Rao (University at Buffalo)
Dr. Stephen Koury (University at Buffalo)
Supported by an NIH Science Education Partnership (SEPA) Award - R25OD10536-1 A1

www.buffalo.edu