Annotation of the *Psedomonas Aeruginosa* Genome LES431 at Locus Tags T223_00095, T223_00100 and T223_00105

Emily Krasinski, Aileen Harvey, Colin Farmer, Alyssa Smith, Shannon McLaughlin, Bridget Hayes, 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 consecutive genes from the microorganism *Pseudomonas Aeruginosa* (T223_00095 – T223_00100 - T223_00105) were annotated using the collaborative genome annotation website GENI-ACT. The Genbank proposed gene product name for each gene was assessed in terms of the general genomic information, amino acid sequence-based similarity data, structure-based evidence from the amino acid sequence, cellular localization data, potential alternative open reading frames, enzymatic function, presence or absence of gene duplication and degradation, the possibility of horizontal gene transfer, and the production of an RNA product. The Genbank proposed gene product name did not differ significantly from the proposed gene annotation for each of the genes in the group and as such, the genes appear to be correctly annotated in the database.

Introduction

Pseudomonas Aeruginosa is a common, gram-negative aerobic, rod shaped bacteria that can cause disease in plants, animals, and humans. *Pseudomonas Aeruginosa* is classified as a Gammaproteobacteria along with other groups of bacteria that are medically, ecologically, and scientifically important, Gram-negative bacteria. it was first discovered in Paris, France in 1882 by Carl Gessard through an experiment that identified the microbe by its water soluble pigments that turned a blue-green when exposed to ultraviolet light. *Pseudomonas Aeruginosa* grows best at 42 degrees Celsius in an aerobic environment, but it is able to reproduce slowly in an anaerobic environment if nitrate is present as a hydrogen acceptor. *Pseudomonas Aeruginosa* is the most common cause of infections of burn injuries and the outer ear. The *Pseudomonas Aeruginosa* bacterium is of interest in medicine because there is a serious risk of it becoming an antibiotic resistant bacteria.
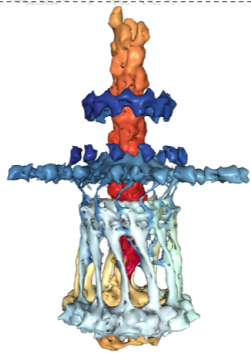

Figure 1. Architectural model of the type IVa pilus machine in a pilated state

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_00095

Peptide deformylase is an enzyme catalyst with ribosome related properties. It is a prokaryotic metalloenzyme, and it could be targeted by an antibacterial known as a peptide deformylase inhibitor. The amino acid sequence matched with its top BLAST hits and GENI-ACT was right in guessing that the sequence was the peptide deformylase enzyme. Peptide deformylase is a protein in *E. coli*, so using the inhibitors could destroy *E. coli* and also help manage the green algae outbursts. I did not find anything different than what GENI_ACT proposed the gene to be. Also, the peptide deformylase is often abbreviated as (PDF). Peptide deformylase is considered to be a biogenesis protein. Peptide deformylase is heavily associated with the organism *Enterobacter cloacae*, a protein in GENI_ACT under the locus_tag T223_00105, the tag near to peptide deformylase, showing that the proteins are very close to each other in terms of location found, gram, organism, and many other features such as they are both related to *Pseudomonas aeruginosa*.
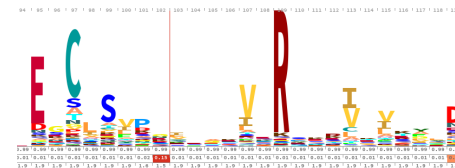

Figure 2. The description of the sequence from Web Logo. It shows key structural residues. Some of these were, E95, C97, R109, H139, D142, and G146
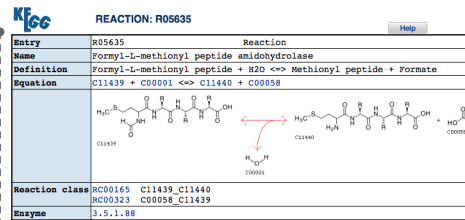

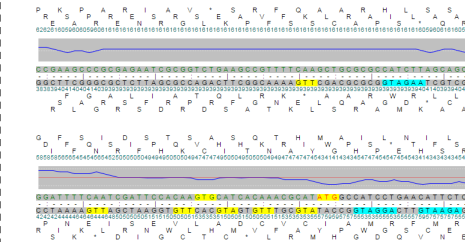Figure 3. This diagram shows the chemical reaction catalyzed by peptide deformylase.

Figure 4. Alternative open reading frame results for T223_00095.

T223_00100

The protein product of this gene was a LysMdomain/BON superfamily protein. This protein showed evidence of a signal peptide being present and no helices within its structure. This protein was not only found in *Pseudomonas aeurginosa* but also in *Enterobacter cloacae*. This bacteria is a rod-shaped, gram negative bacteria. Although this is not a normal pathogen for humans, it has been known to be associated with urinary tract and respiratory tract infections. This is a protein highly related to *Pseudomonas aeruginosa* Peptide Deformylase. The Signal IP results indicated that this protein is most likely a signal peptide.
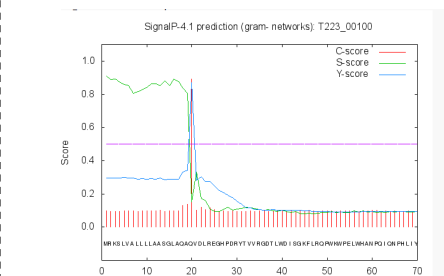

Figure 5. Signal IP showing the presence of a signal peptide for T223_00100.

T223_00105

The initial proposed product of this gene was a DNA processing protein that resides inside of the bacteria, where it will stay. The TMHMM results predicted that there are no transmembrane helices. My research told me that the protein may have multiple locations in the cells and is involved in DNA processing/protection. The protein will not leave the bacteria so it can do it's job – process/protect DNA
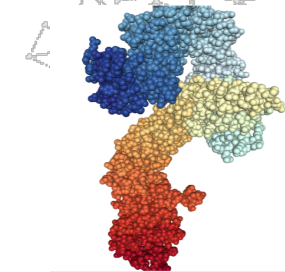

Figure 6. A model of a DNA processing protein

Conclusion

The GENI-ACT proposed gene product did not differ significantly from the proposed gene annotation for each of the genes in the group and as such, the genes appear to be correctly annotated by the computer database.

Gene Locus	Geni-Act Gene Products	Proposed Annotation	Functions
T223_00095	Peptide deformylase	Peptide deformylase	metalloenzymes essential for cell growth
T223_00100	LysM domain/BON super family protein	LysM domain/BON super family protein	Aids in the formation of the bacterial cell wall
T2233_00105	DNA processing protein DprA	DNA Protecting Protein DrpA	Processes DNA- stays inside bacteria

References

Gammaproteobacteria. Retrieved March 18th, 2017, from https://www.boundless.com/microbiology/textbooks/boundless-microbiology-textbook/microbial-evolution-phylogeny-and-diversity-8/proteobacteria-103/gammaproteobacteria-543-4726/
LysM Domain. Retrieved March 18th, 2017, from https://www.ebi.ac.uk/interpro/entry/IPR018392

Acknowledgments