Annotation of the Bacillus anthracis from Locus Tags BA_0633 to BA 0635

Taylor Clark, Arianna Galetta, Amalia McKenna, Marissa Nugent, Ryan Zimmerman, Sean Zimmerman, Richard Salton Gates Chili High School, Rochester New York and the Western New York Genetics in Research and Health Care Partnership

University at Buffalo

SCIENCE EDUCATION rted by the National Institutes of Hea

> **Gates** Chili Central

Abstract

A group of 3 genes from Bacillus Anthracis (Bant 650002 to Bant 653696) were annotated using the Genomics Education National Initiative - Genome Annotations Toolkit (GENI-ACT). GENI-ACT was used to examine different functions of a specific protein. This was done through completing specific tests on the GENI-ACT program that gave insight to the amino acid sequence and the way it sets up the polypeptides in the gene. Certain programs that were used include T-Coffee, P-fam, and CDD that helped pinpoint these functions with the DNA tag. Bant 0633 was tested and appeared to be a gene that deals with spore germination. Bant 0634 and Bant 0635 were also examined and were mainly affiliated with spore germination as well. The KEGG program discovered that the gene had no enzymatic function related to it.

Introduction

Bacillus anthracis was first recorded in 1491. The first breakthrough was made by Pierre Rayer and Casimir-Joseph Davaine. It originates from an agent of anthrax, a common disease that occurs in mainly livestock and sometime humans. This is a rod shaped, gram positive, endospore forming bacterium. Bacillus anthracis can be grown in an aerobic or anaerobic environment. The endospore has a very thick cell membrane with many layers and can remain inactive for years, and then once again become active once it has been placed in a favorable environment.

Gram staining is a process used to see if a bacterium is gram positive or gram negative, also showing the chemical makeup of the cell wall. Gram-staining test for the thickness of the peptidoglycan in the cell wall, which could make the protein hard to eradicate.

An endospore stain can be used to visualize bacterial endospores. Endospores are formed by a few genera of bacteria, such as Bacillus. The endospore stain for Bacillus Anthracis

Geni-act was used to help identify functions of various loci on a genome that we were given. Different programs are used to pinpoint the exact process of specific locations on this genome. When we figure out the functions of each loci, we can come closer to figuring out the main function of the protein



Figure 1 - Picture of the morphology of Bacillus anthracis

Methods

Modules of the GENI-ACT (http://www.geni-act.org/) were used to complete Bacillus Anthracis 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?
Module 6- Enzymatic Function	KEGG, MetaCyc, E.C. Number,	In what process does my protein take part?

Results

BA 0633 to BA 0635 The initial proposed products of all of these genes by GENI-ACT related to spore germination. These gene products were supported by the top BLAST hits for the amino acid sequences, the presence of a signal peptide within the amino acid sequence of BA 0633, the cytoplasmic membrane localization of BA 0635 and the cellular location of all three amino acids using the pSort tool. During CDD all three gene groups did not have any COG number and name, nor any enzymatic function in Kegg. Finally, the proposed genes are all related to spore germination of



MTEQKEQQESTGOKHMFPSLPENINYIENKLCHSDDI KKLDLPFQNGKGTILYIESLADPNLIHQLALEPLLTRSDL SLDKAFATLNMKKETNLNYGVQLLLQGKSLYFHEHVDSFC * IFETAJSLKRDIAEPDNEGIVRGPHTGFVEDLATNLISIR * KLIKSPNIVVKYFTLGEEMHTKVAIAYMQNIANDDLVTEV * KRRLETIKTDALMPPGYIQEFIEDTSFSPFPQQLNTERPD * RVAANLMEGRVAIFSDGDPTALIVPVTLFAFYQSPDDYNN ***** RWIVGSEVRMIRLVSELIAELLPAIYIATVAEHPDVLPLE ***** LVYTIKASLEKVPLPPIFEALLMELIFELLREAGIRLPSR VGQTIGIVGGLVIGDAIVKAGLVSYTMIIVVALTAISSFL * VPSNDMSSAVRILRFPLMLLAALFGYVGISFGLIITFVHL COLHSFHTPYLSPVAPMRIKDMKDSFVRLPIWSFWERPHD PKPKKMORQHVTREEEDGDKHAK

Figure 3 - BA 0635 WebLOGO for spore germination. The logo shows the significant presence of highly-conserved areas of amino acids in the protein. This means this given protein sequence and structure can evolve, function, and exist independently of the rest of the protein chain.

Putative conserved domains have been detected, click on the image below for detailed results.

IIPTU SPD.		9 1	19	151	20		30 	
pecific hits	(spore_ger_x_C						
uperfamilies				Spore_Ge	erAC superfan	ily		

Figure 4 - BA 0633 conserved domains which are found using the amino acid sequence.

Putative conserved domains have been detected, click on the image below for detailed results.

9	<u>5</u> 19 25 99 17 49					
query seq, 🛛 🗰 Specific hits	GerA					
Superfamilies	GerA superfamily					

earch



Conclusion

The GENI-ACT proposed gene's main function relates to spore germination.

		Sector Contraction of the sector of the sect
Gene Locus	Geni-Act Products	Proposed Annotation
0633	Spore germination protein GerKC	Spore germination
0634	Spore germination protein GerKB	Spore germination
0635	Spore germination protein KA	Spore germination

References

Sims et al. (2017). Complete genome sequence of Bacillus Anthracis type strain (541T). Standards Genomic Sciences, 12 - 20.

Acknowledgments

Supported by an NIH Science Education Partnership Award (SEPA) - R250D010536

www.buffalo.edu