Annotation of the *Borrelia burgdorferi* Genome at Locus Tags BB_0149 and BB_0151

Hanna Tormey, Claire Volk, Abigail Tarun, and Kate Carretto
Alfred Almond Central School, Almond, NY and The Western New York Genetics in Research and Health Care Partnership

SEPA SCIENCE EDUCATION PARTNERSHIP AWARD
Supported by the National Institutes of Health
ALFRED-ALMOND CENTRAL SCHOOL

Abstract

Two genes from the organism, *Borrelia burgdorferi*, were annotated using the website, GENI-ACT. The locus tags of these two genes are BB_0149 and BB_0151. This microorganism is a pathogen that causes Lyme Disease. We utilized the GENI-ACT website to organize our data and results from other sources. We used the Basic Information, Sequence-Based Similarity, Structure-Based Similarity, Cellular Localization, and Open Reading Frame modules within the website to optimize our research. Other sources of information, such as BLAST and WebLogo helped us to determine structure and function of the protein that the gene codes for. These research databases helped us to determine that the BB_0149 gene coded for a flagellar hook protein while the BB_0151 protein codes for the protein N-acetylglucosamine-6-phosphate deacetylase or nagA.

Introduction

Borrelia burgdorferi is a spirochete, tick-borne obligate parasite. The genome of this pathogen includes a linear chromosome as well as a large number of smaller DNA molecules or plasmids. Some of these plasmids are linear while others are circular. This type genomic format is atypical in the bacterial world, but the majority of the genes on these chromosomes are common in bacteria. One additional characteristic of the *Borrelia burgdorferi* is its short lifespan, which is reflected through its small genome which lacks the conventionally recognizable machinery for synthesizing nucleotides, amino acids, fatty acids, and enzyme cofactors.

Borrelia burgdorferi is a parasite that causes Lyme Disease in mammal hosts. Lyme Disease is transmitted to humans among other mammals through the bite of a blacklegged tick that has become infected with the parasite. Some symptoms of this disease include fever, headache, erythema migrans, and fatigue. The OspC gene on the *Borrelia burgdorferi* chromosome is responsible for allowing the parasite to survive the initial immune response from the host. The VIsE gene allows the parasite to continue to fight the immune response from the host, apart from the initial response. Lyme Disease has become an epidemic in Western New York, so study and genomic mapping of the parasite that causes it is needed to create a medical response to the parasitic attack.

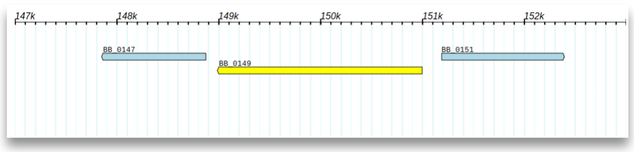

Figure 1 - Locus Tags The locus tags and relative position of the genes under investigation in this research

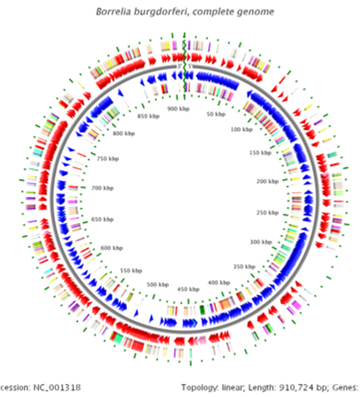

Figure 2 – A map of the complete genome of *Borrelia burgdorferi*.

Methods

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

Modules	Activities	Questions Investigated
Basic Information	DNA Coordinates and Sequence, Protein Sequence	What is the sequence of the gene and protein? Where is it located in the genome?
Sequence-Based Similarity	Blast, CDD, T-Coffee, WebLogo	How similar is the protein under investigation to other proteins in GenBank?
Structure-Based Similarity	TIGRfam, Pfam, PDB	What functional domains are present in the protein under investigation?
Cellular Localization	Gram Stain, TMHMM, SignalP, LipoP, Psortb, Phobius	Is the protein under investigation located in the cytoplasm, secreted, in the periplasm or embedded in the cell membrane or cell wall?
Alternative Open Reading Frame	IG/M, ORF, Blast	Is there evidence of an alternative starting codon for the protein and if so, what are its hypothetical DNA coordinates?

Results

BB_0149: Flagellar Hook-Associated Protein FliD

The computer pipeline proposed product of this gene was a flagellar hook associated protein FliD, a small, highly curved tubular structure located in the periplasm that connects the flagellar motor to the long helical propeller of the flagella. This gene product proposal was supported by the top BLAST hits for the amino acid sequence, the presence of highly-conserved functional domains in the CDD results, the periplasmic score in PSORT-B, and the correctly proposed DNA coordinates from ORF.

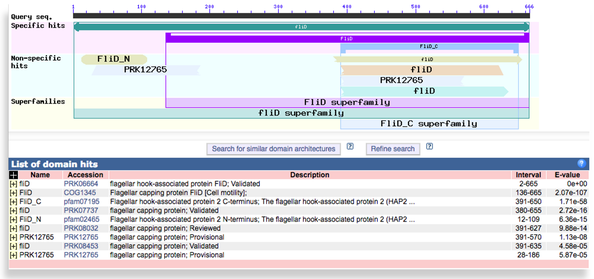

Figure 3 – BB_0149 CDD Results a well-conserved list of functional domains and families specific to the flagellar hook protein FliD, most notable were the Pfam results, flagellar hook associated protein 2 N and C terminus which had several key structural residues that were nearly identical amino acid sequences to BB_0149

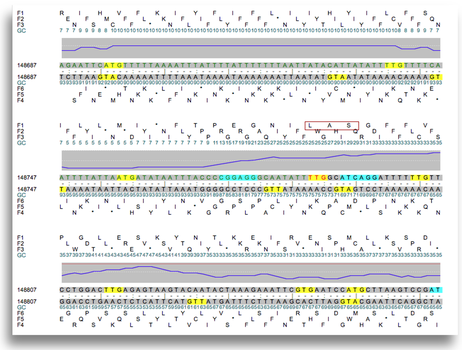

Figure 4 – Graphic Sequence Viewer Results from ORF correctly proposed the starting DNA coordinates as shown by a nearby Shine-Dalgarno sequence (blue) and the corresponding starting codon LAS on the top (red border).

BB_0151: N-acetylglucosamine-6-phosphate deacetylase

This codes for the enzyme N-acetylglucosamine-6-phosphate deacetylase (NagA) which catalyzes the deacetylation of N-acetylglucosamine-6-phosphate (GlcNAc-6-P) to glucosamine-6-phosphate (GlcN-6-P), which is contained in the cytoplasm. This deacetylation provides energy for the cell or is used in order to maintain the peptidoglycan wall. The prevalent NagA BLAST hits, presence of the Amidohydrolase family in the Pfam records, and cytoplasmic score in PSORT-B all support these claims.

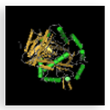

Figure 5 - BB_0151 The TIGR00221 diagram of Central intermediary metabolism from the CDD results in the BLAST database.

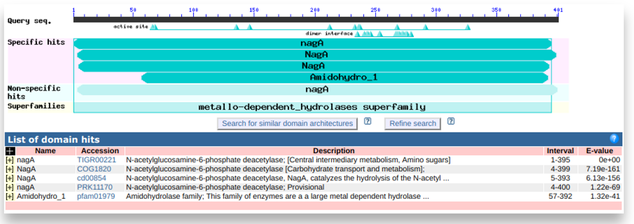

Figure 6 - BB_0151 CDD Results limited domain hits for TIGRFAM and Pfam in the nr BLAST database showing the deacetylation of NagA.

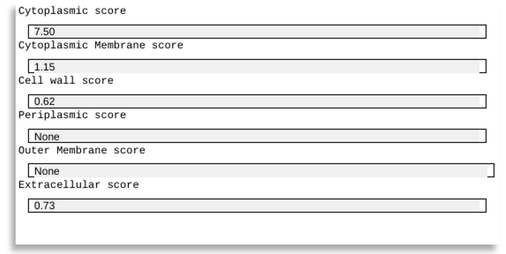

Figure 7 - BB_0151 PSORT-B scores from the Cellular Localization Module showing the location of NagA in the cytoplasm.

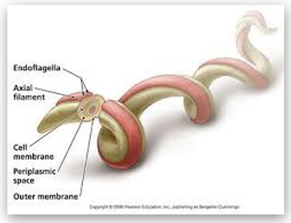

Figure 8 - Model of the *Borrelia burgdorferi* highlighting the spirochete structure and flagella.

Conclusion

Our GENI-ACT proposed gene products were the same as the proposed gene annotation. We have no changes to propose. We believe that the computer correctly annotated the BB_0149 and BB_0151 gene on the *Borrelia burgdorferi*. We found that the BB_0149 gene codes for a Flagellar Hook Associated Protein FliD and the BB_0151 gene codes for N-acetylglucosamine-6-phosphate deacetylase.

Locus Tag	Pipeline Annotation Product Name	Proposed Annotation	Changes Proposed?
BB_0149	Flagellar Hook Associated Protein FliD	Flagellar Hook Associated Protein	No
BB_0151	N-acetylglucosamine-6-phosphate deacetylase	N-acetylglucosamine-6-phosphate deacetylase	No

References

Tilly, K., Rosa, P. A., & Stewart, P. E. (2008, June). Biology of infection with *Borrelia burgdorferi*. Retrieved May 13, 2019, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2440571

University at Buffalo The State University of New York
Jacobs School of Medicine and Biomedical Sciences
buffalo.edu

Acknowledgments

Dr. Rama Dey-Rao & Dr. Stephen Koury
Supported by an NIH Science Education Partnership (SEPA) Award - R25GM129209