Annotation of the *Brucella abortus* Genome From DNA Cordinates 1818906 to 1821062 and 3606 to 3830

Ashley Ameis, Danielle Herrick, Elizabeth Miller, and Mr. Kirk Winters
Pioneer Central High School and The Western New York Genetics in Research Partnership

Abstract

A selection of two genes from the organism *Brucella abortus* (BAA13334_I00004, BAA1334_IO3018) were assigned for further investigation. Through GENI-ACT an investigation was conducted to find information on each genome annotation. We investigated general genomic information and the amino acid sequence-based similarity of data from the databases. Investigation was also conducted to find the cellular localization, we also used the amino acid sequence to find structured-based evidence. We entered our gene sequences from the gene bank into the data bases and most were successfully identified, although not all genes were successfully matched with the database. It was through the GENI-ACT databases that investigations were conducted on the gene sequences from the organism *Brucella abortus*.

Introduction

Brucella abortus is a microscopic, gram negative coccobacilli. This bacterium is the cause of Brucellosis in cattle, also know as Bangs disease. It is zoonotic, which means that it can be transmitted between mammals. These bacteria are able to get inside of a mammal through cuts or the mucous membranes. The phagocytes are then infected and travel through the blood stream to the lymph nodes. Then the lymph nodes release the bacteria back into the blood stream where it can travel throughout the body. This bacteria causes headaches, weight loss, fever, and muscle pain. The major effect of this bacteria occurs in pregnant cows and causes abortions. The abortions are caused by the infection of the placenta, which leads the cow to expel the fetus. This disease is spread worldwide, but the areas where it is the most prevalent are Africa, India, Mexico and South America. There is a medication made from a combination of Tetracycline and Streptomycin. They are prescribed this medication for three to six weeks. There is a vaccination that is widely used on animals, but does not work on humans. There are genetically engineered vaccinations that use recombinant genes that are put into modified Brucella abortus strains, which are then used as vectors to help the vaccinated mammal to build up an immunity against this disease. Due to this disease having the greatest effect on cattle, they are the ones that are the biggest market for these vaccines.

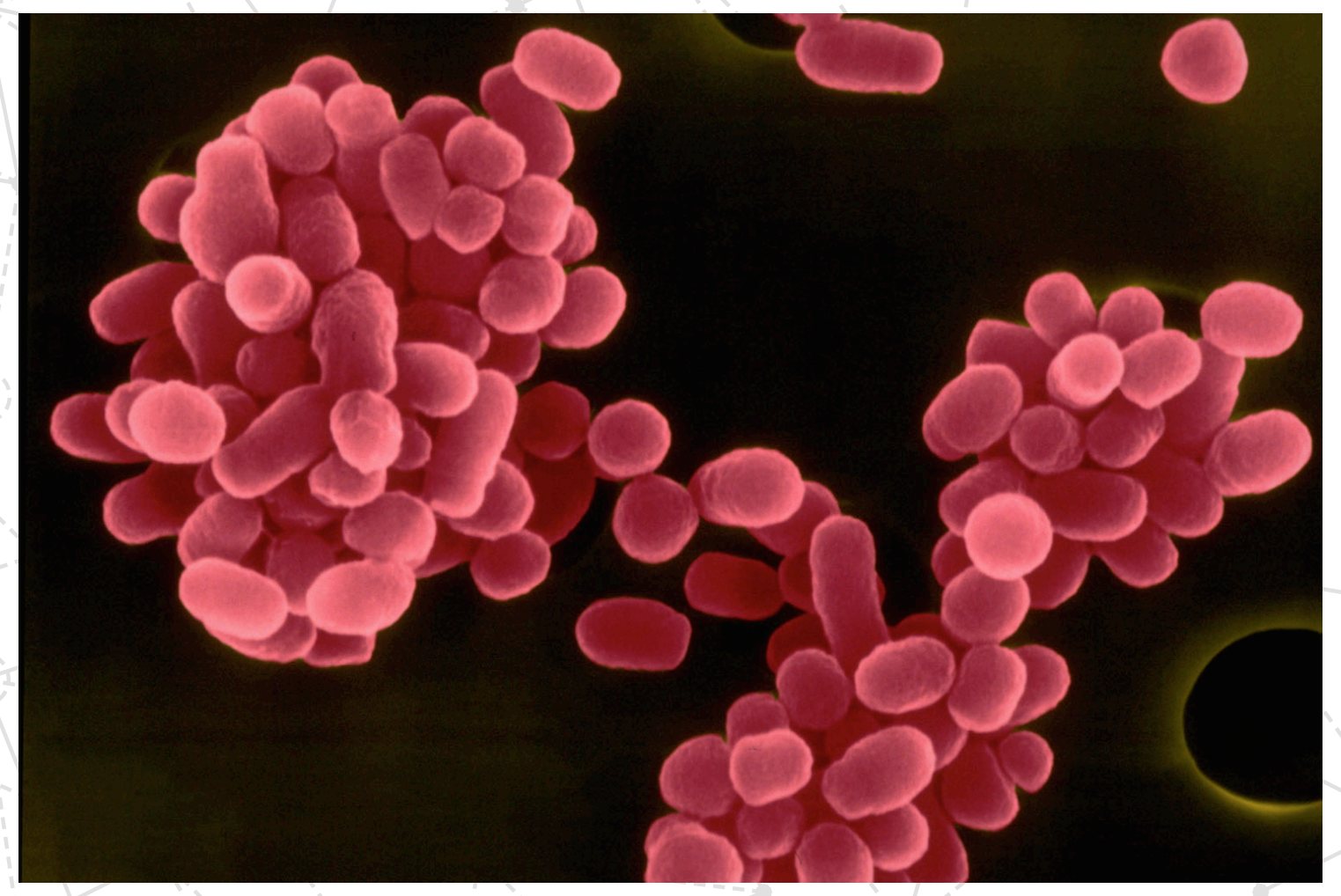

Figure 1. A picture of the *Brucella abortus* bacteria under a electron microscope (City News)

Methods

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

Methods	Actvities	Questions Investigated
Module 1-Basic Information Module	DNA Coordinates and Sequence, Protein Sequence	What is the sequence of my gene and protein? Where is my gene located?
Module 2- Sequence Based Similarity Data	Blast, T-Coffee, CDD, WebLogo	How does my amino acid sequence compare with other sequences that have been discovered?
Module 3- Cellular Localization Data	Gram Stain, SignalP, PSORT, TMHMM, Phobus	Where is my protein located in the cell? Is it secreted or embedded in the membrane?
Module 4- Structure Based Evidence	TIGRfam, Pfam, PDB	Are there any functional domains that are located in my protein?

Results

BAA13334_I00004

Through GENI-ACT gene sequence BAA13334_I00004 was given for further investigation. As this gene sequence was further explored, there were very few results or matches through the data bases. This may be because of a very specific and unique makeup that isn't compatible with many other sequences. Through this exploration of gene sequence BAA13334_I00004 it could be concluded that this gene sequence is located within the many cell membrane's. This can be concluded based on PSORT-B and Phobius. Both of which are databases through GENI-ACT which are used to compare sequences. Results from both these comparisons showed a strong representation for a membrane location, almost completly ruling out a cytoplasmic sequence. Through these GENI-ACT database many discoveries were able to be made about this gene sequence.

BAA1334_103018

The gene segment from BAA1334_103018, which comes from the organism *Brucella abortus*, was thoroughly researched through the GENI-ACT database. Within this segment, there is a possibility for there to be a polypeptide, but there were no results for a transmembrane helixe. The results for the T-COFFEE and Multiple sequence alignment were extremely long because the amino acid sequence was 718 units long. This resulted in a WebLogo and HMM Logo that was also at considerable length. Due to the large amount of results, we were able to conclude that this sequence can be found in a wide variety of organisms. Through the GENI-ACT database we were able to conclude that this specific sequence is of considerable length and is very wide spread.

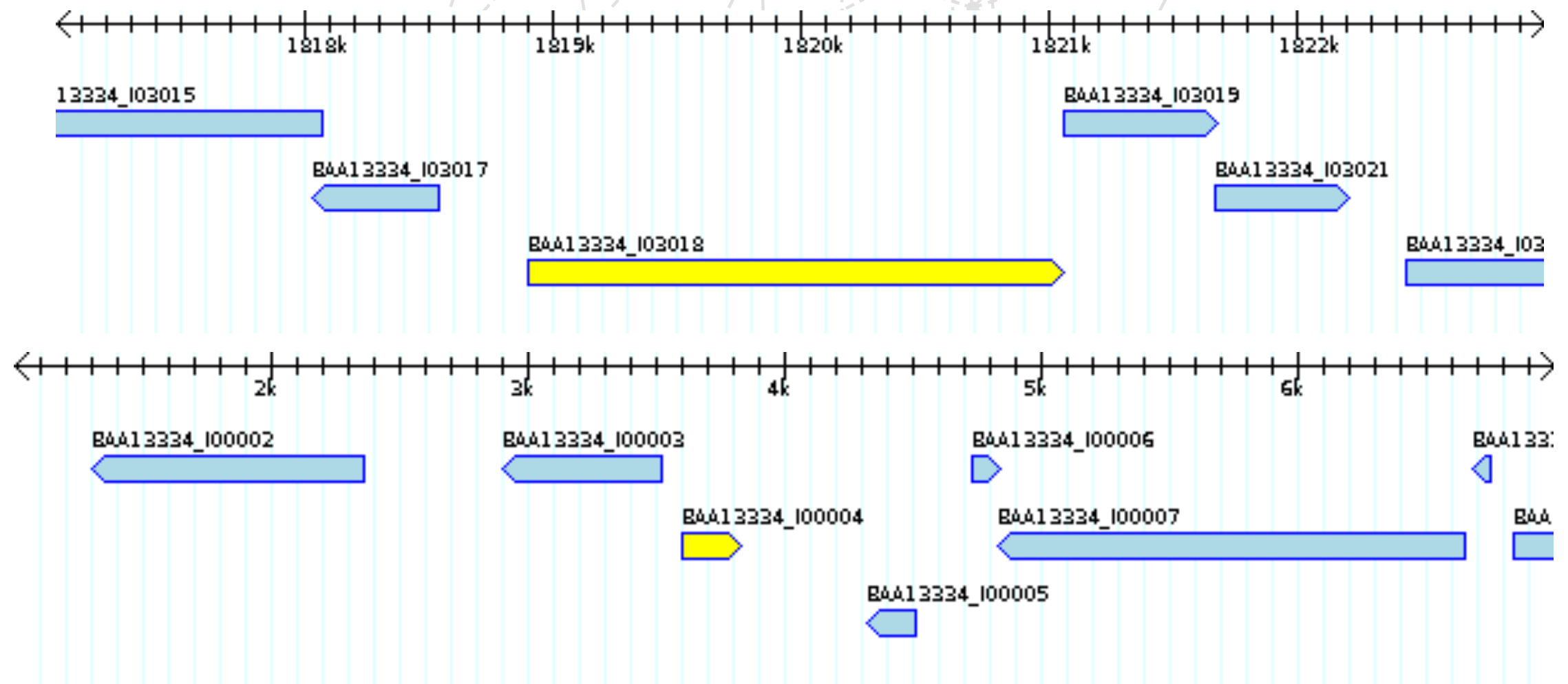

Figure 2 – The yellow arrow represent the gene segment. The arrow is pointing to the right which indicates that this gene segment is on the top of the DNA strand.

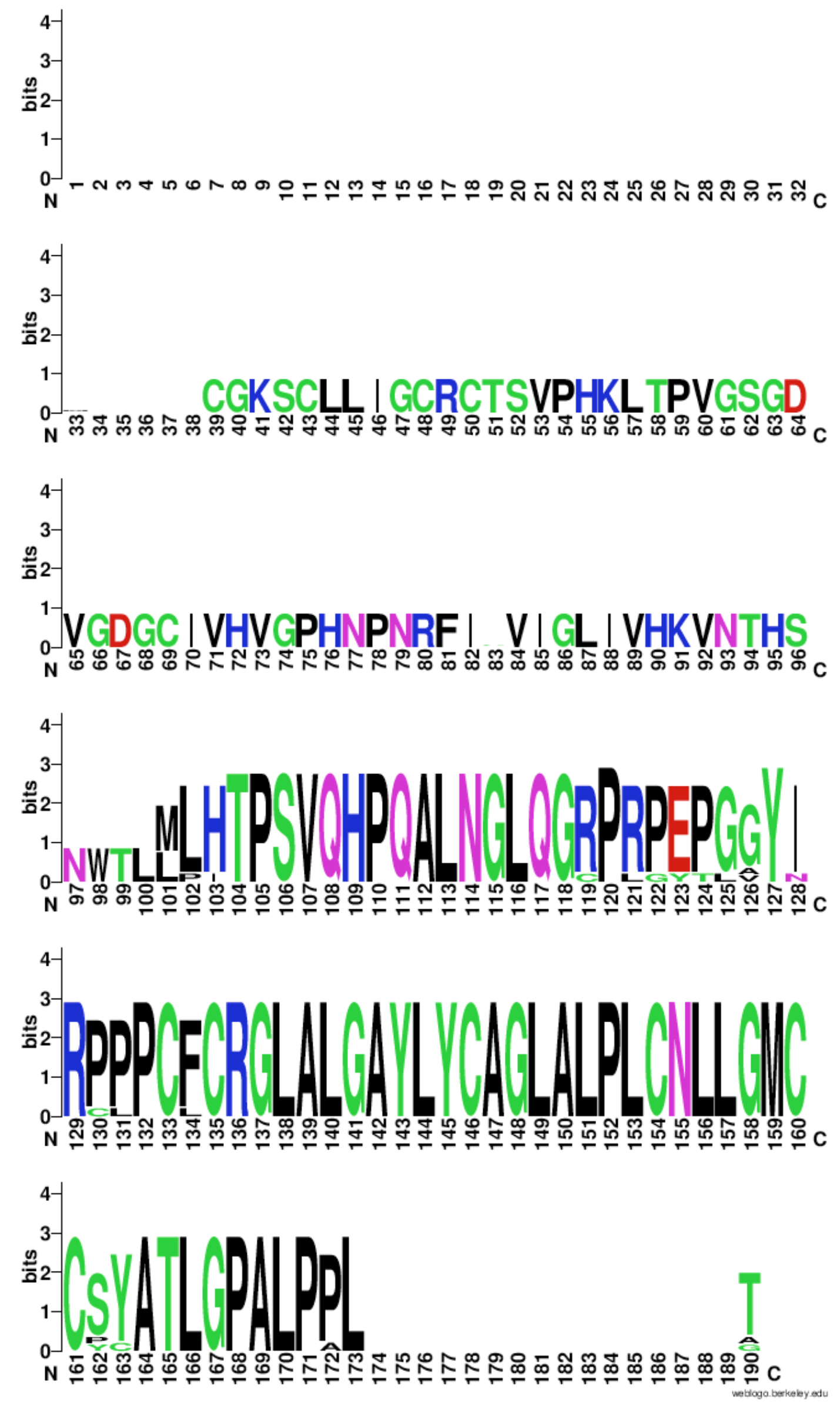

Figure 3 – There is very little consistency in the amino terminal end of the alignment, but more consistency in the C-terminus with hydrophobic (black) and polar (green) amino acids predominating which leads to the conclusion of having more reliable sequence conservation near the end compared to the beginning

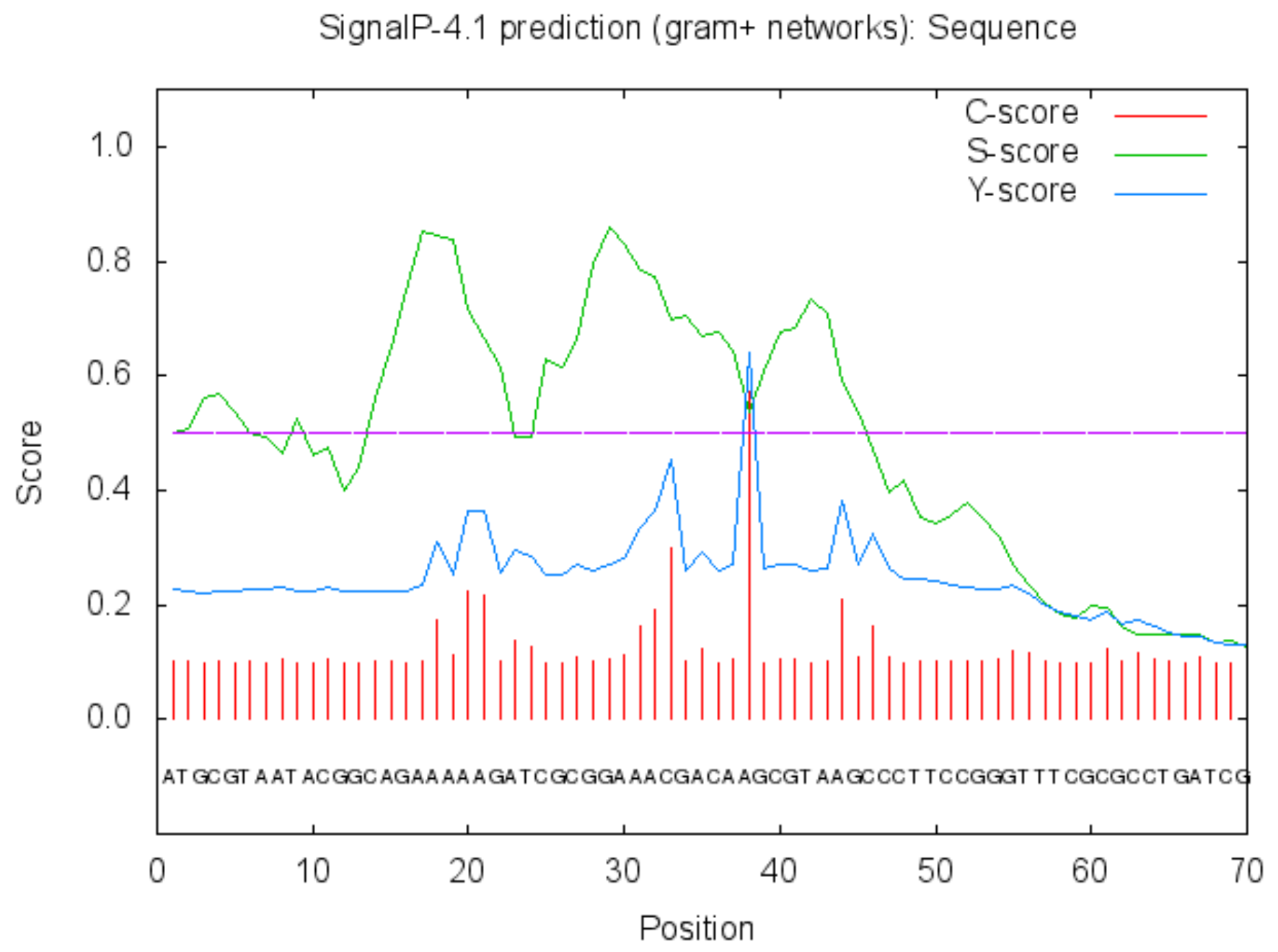

Figure 4 – In this graph the data cleaves between the amino acids 37 and 38 which shows where the single peptide is located in the gene sequence.

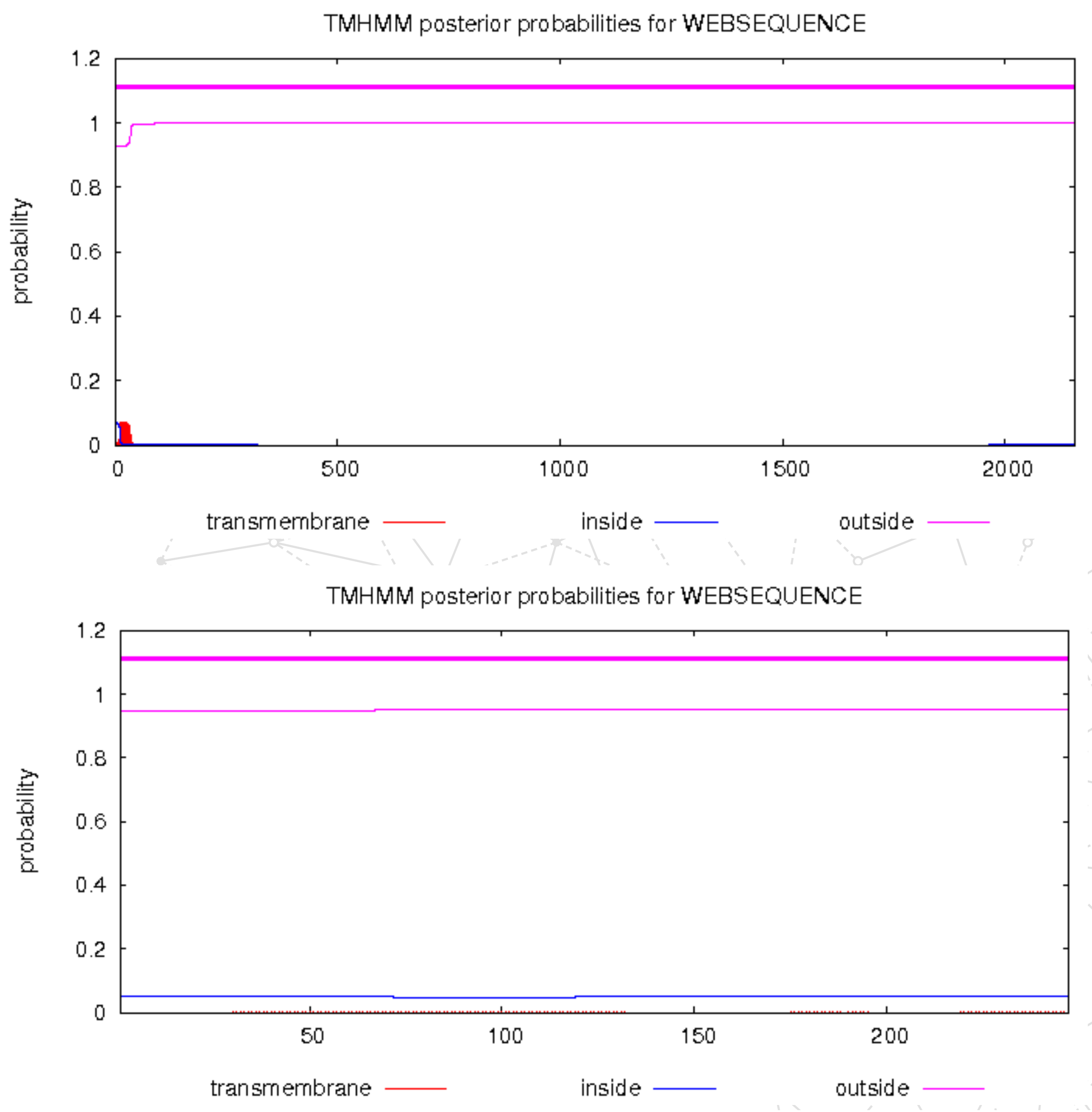

Figure 5- These graphs demonstrates that there are no transmembrane helixes within the cytoplasmic membrane.

Conclusions

Through GENI-ACT and its surrounding databases, these gene sequences were able to be explored with the goal of finding more information on the gene sequence. This information revolved around the discovery of the function and exact location of the gene. Through GENI-ACT and the surrounding data bases, results were able to be collected, although some were inconclusive. With these results we were able to investigate and draw conclusions involving the gene sequences. Further research could be done to further explore the full capability of these genes.

References

http://www.citynews.ca/wp-content/blogs.dir/sites/10/2013/04/brucella.jpg

Acknowledgments

Supported by an NSF Innovative Technology Experiences for Students and Teachers (ITEST) Award - 1311902

www.buffalo.edu