Annotation of the *Clostridium perfringens* Genome at Locus Tags CPF_1661, CPF_1668 and CPF_1675

Darby Berdine, Fox Putrino, Bhuiya Faisal, Mrs. Kelly Cal and Mr. Todd Richards City Honors School, Buffalo, NY and the Western New York Genetics in Research and Health Care Partnership



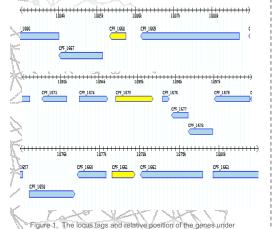


Abstract

The microorganism Clostridium perfringens (CPF_1675 - CPF_1661) was 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, and potential alternative open reading frames. 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 by in the database. We found it interesting that the bacteria turned out to be a food based illness. Locus tag CPF_1661 or Flaverodoxin is most likely found in the Cytoplasmic Membrane.

Introduction

Clostridium perfringens is a rod-shaped food borne illness that is associated with undercooked meats. It's a bacteria that produces harmful toxins to humans. Food caused by Clostridium perfringens is usually mistaken for the common flu. The bacteria is found in food ranging from temperatures of 70-140 degrees fahrenheit. Clostridium perfringens can be diagnosed by Nagler's reaction where the suspect organism is cultured on an egg yolk media plate. One side of the plate is anti-alphatoxin, the other side does not. This bacteria is usually found in contaminated meats, stews, soups, and gravies.



on in this research (CPF 1668, CPF 1675, and CPF 1661)

Methods

Modules of the GENI-ACT (http://www.geni-act.org/) were used to complete *Clostidium pertringens* 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

CPF 1675:

The initial proposed product of this gene by GENI-ACT was a Uncharacterized aminopeptidase Mb1368. This came from the species Mycobacterium bovis with a length of 344 Amino Acids. This proposed product didn't have very much information on it because of its missing characterization. The discovered Nucleotide sequence turned out to be:

ARTHMAN GOOD TO CONTROL OF THE ARTHMAN CREEKES CREE

and Amino acid sequence:

INDIA
HTKIKITDIDGFKLGHAQDFEGATQCTVLLCEEGASGQVDVRQQAPGTRETDLLNPMEMVDKVI
VVLSQGSAFQLDSCSGVMEYLENKNVGFDVQVAKVPIVCGAVLFDLACGNPKVRPNKEMGLE
KNSETVLDSKNGNICCGTQATVGKALNGKLANKGGPSVAVVOGDKGVAIVQNNSLGOVDI
NNNNKIJAGJSQSOPINSFINEESLLANYSPRIVFKGNTTIGCIVTKNQGAIVGNNSLGOVDI
NNNNKIJAGJSQSOPINSFINEESLLANYSPRIVFKGNTTIGCIVTKNQGFKKAGNAKKISIMAO

CPF 1661:

The initial proposed product of this gene by GENI-ACT was a Flaverodoxin. Flaverodoxin is an electron transfer protein, that comes from the organism Desulfovibrio gigas. This gene product proposal was supported by the top BLAST hits for the amino acid sequence, the presence of well-curated functional domains within the amino acid sequence, the cellular location of the amino acid sequence, and the enzymatic function of the amino acid sequence.

Nucleotide Sequence:

Amino Acid Sequence:

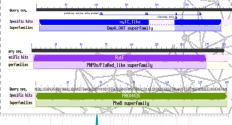
KECVINLPCRDMVKFVDFCGVKSGRKVDKIKHFGYALDKGIAVSTPSIEKAPVAI GTHDWFLFEVLNVKVDESLIDEKGKIHFNKANLLAYSHGEYFGLNSKPLGSFGYSIKKKKKSR

CPF 1668:

The initial proposed product of these genes by GENI-ACT was an ACT domain protein. This gene product proposal was supported by the top BLAST hits for the amino acid sequences, the presence of well-curated functional domains within the amino acid sequences, the cellular location of the amino acid sequences, and the enzymatic function of the amino acid sequences.

The Nucleotide sequence was:

GATAACAGAAGCTGCAAAGCAGCTGGTATTAGTAGAAGTGTTTATTATAAGTATAAAGACTATATATTGAATTTGAATTTGAATTTGAAACCCTTCAAGGTAGAAAA GTTATTTAATATGGTGTTACCCATGAAAAGGGAGTTTTAATTGACTGTTTAATATTATTCAAGTTGGGGGAAATATACTTACAATTAGACCAGG GCTACCTATACATGGATTAGCTCAAGTATGAACTTACACATAGACCAATCAAGTGAAGTGAATTAAAGGGAAATTAACTTACAATTAGAATTA



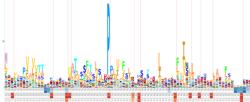
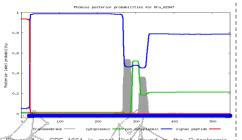




Figure 2. CPF_1661 HMM logo

Figure 3. CPF_1675, Web sequence logo



Floure 4 CPF_1661 is most likely found in the Cytoplasmic Membrane, based on the results of the Phobius probability graph and there was PSORT-B score of 9.55.

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.

	7	The state of the s
Gene Locus:	Geni-act Gene	Proposed Annotation:
	Products:	411
CPF_1661	Flaverodoxin	Flaverodoxin
-	12-11	
CPF_1675	Uncharacterized	Uncharacterized
- and and a state of the state	aminopeptidase Mb1368	aminopeptidase Mb1368
CPF_1668	Act domain protein	Act domain protein
	\ \	1 1 381]

References

Sims et al. (2009). Complete genome sequence of Clostridium Perfringens type strain (541T). Standards Genomic Sciences,12 - 20.

Acknowledgments

Supported by an NIH Science Education Partnership (SEPA) Award - R25ODO10536

www.buffalo.edu