Annotation of the *Campylobacter jejuni* Genome from DNA Coordinates from locus tags CJE0008 to CJE0010

Janette Tatum, Wahida Jannat, Mi Rasa and Dana Pryor Moncrieffe
Frederick Law Olmsted School and the Western New York Genetics in Research and Health Care Partnership

Abstract

Students in the research group of Frederick Law Olmsted’s Science Club have been assigned to Three consecutive locus tags to annotate in the organism, *Campylobacter jejuni*. The genome annotation website GENI-ACT allows us to research basic information, sequence based similarity data, cellular localization data, alternative open reading frame, and structure based evidence. Interpretations of data are depicted by information given from BLAST results, using the amino acid sequence of *Campylobacter jejuni*. A hypothesis of whether the predicted product name is well suited for the gene is found at the very end of the research. According to the overall data of this group, our findings did not contradict the Genbank product names of the genes, except for Locus Tag CJE0010.

Introduction

*Campylobacter jejuni* was noticed in 1886 by Theodor Escherich in kitten’s stool. The studies that progressed after that tells us that researchers discovered that *Campylobacter jejuni* was present in untreated and filtered lake water, unpasteurized milk, etc. It is known as the leading cause of bacterial food-borne diarrheal disease throughout the world. *Campylobacter jejuni* was also known as the most frequent antecedent to a form of neurovascular paralysis known as Guillain-Barré syndrome in 2000.

Current studies tell us that *Campylobacter jejuni* is currently the leading cause of bacterial gastroenteritis in humans, that disease in humans is mainly caused by the azonic pathogen *Campylobacter jejuni* and that chickens and poultry meat are common hosts for the bacteria. *Campylobacter jejuni* targets C-type lectin (SIGNR1, SIGNR3) and immunoglobulin-like receptors (TREM2, TREM3, LMR3, LMR9) causing gastroenteritis and Guillain-Barré syndrome.

Our team annotated three different gene sequences from the *Campylobacter jejuni* genome. After analyzing the sequences, we have determined the gene products: a small sub unit of glutamate synthase and ribonuclease HI to have been named correctly by the GENI-ACT.

Methods and Materials

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

- **Module 1: Basic Information:**
  - **Activities:** DNA Coordinates and Sequence, Protein Sequence
  - **Questions Investigated:** What is the sequence of my gene and protein? Where is it located in the genome?

- **Module 2: Sequence-Based Similarity Data:**
  - **Activities:** Blast, CDD, T-Coffee, GenLog
  - **Questions Investigated:** Is your sequence similar to other sequences in GenBank?

- **Module 3: Cellular Localization Data:**
  - **Activities:** Gram Stain, TRANSMEM, SignalP, PSORT, ProRule
  - **Questions Investigated:** Is my gene in the cytoplasm, membrane, or extracellular?

- **Module 4: Alternative Open Reading Frame:**
  - **Activities:** bNG Sequence Viewer For Alternate ORF Search
  - **Questions Investigated:** Are there functional domains in my protein?

- **Module 5: Structure-Based Evidence:**
  - **Activities:** TIGRam, Phan, PSORT
  - **Questions Investigated:** Do the amino acid sequence of my protein been called correctly by the computer?

Results

*Campylobacter jejuni* CJE0008:

The initial proposed product of this gene by GENI-ACT was a small subunit of glutamate synthase. Numerous BLAST hits determined the gene product to be the same as the initial gene product predicted by GENI-ACT, using the amino acid sequence, the presence of well-conserved functional domains within the amino acid sequence, the cellular location of the amino acid sequence, and the enzymatic function of the amino acid sequence. Therefore final annotation is a small subunit of glutamate synthase.

*Campylobacter jejuni* CJE0009:

GENI-ACT initially proposed the product of this gene to be ribonuclease HI. The top BLAST hits for the amino acid sequence, the presence of well-conserved protein functional domains within the amino acid sequence, the trans membrane topography of the amino acid sequence, and the cellular location of the amino acid sequence supported these proposals. In conclusion, the gene product is ribonuclease HI. Ribonuclease HI is involved in the degradation of the ribonucleic moiety on RNA-DNA hybrid molecules carrying out endonucleolytic cleavage to 5-phospho-monoester (5).

*Campylobacter jejuni* CJE0010:

The initial proposed product of the gene predicted by GENI-ACT was competence protein ComEA. The top BLAST hits for the amino acid sequence were multispecies DNA-binding protein. The top COG hit was COG1555: ComEA, with the description of DNA uptake protein ComE and related DNA-binding proteins. Competence is the ability of a cell to take up exogenous DNA from its environment, resulting in transformation. It is widespread among bacteria and is probably an important mechanism for the horizontal transfer of genes. ComEA has been shown to be an integral membrane protein, as predicted from hydrophathy analysis, with its C-terminal domain outside the cytoplasmic membrane (6). TMHMM predicts CJE0010 to have one transmembrane helix (Figure IV) and Psort predicts CJE0010 to be a membrane protein (Figure V) consistent with ComEA being an integral membrane protein.

Conclusion

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

References

1. https://zoonooticology.files.wordpress.com/2013/05/campylobacter-e1367919852939.jpg

Acknowledgments

Supported by an NIH Science Education Partnership (SEPA) Award - R25OD10536. Dr. Stephen Koury. Dr. Rama Dey-Rao