Annotation of the *Campylobacter jejuni* strain RM 1221 at Locus Tags CJE0498, CJE0699, CJE0400, CJE0415, CJE0820 and CJE0188

Randsley Berry, Brianna Bubernak, Michele Lin, Karly Sondericker, Lacey Standish, Alyssa Wright ,Laura O'Donnell and Elizabeth Winters
Attica High School and The Western New York Genetics in Research Partnership

Abstract

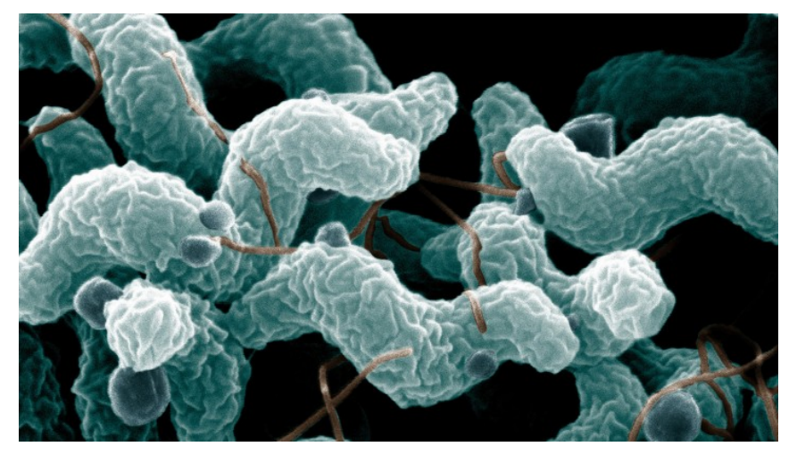
A group of 6 genes from the microorganism *Campylobacter jejuni* were annotated using the collaborative genome annotation website GENI-ACT. The Genbank proposed gene product name for each gene was assessed in terms of the general genomic information, amino acid sequence-based similarity data, structure-based evidence from the amino acid sequence, and cellular localization data. 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 the computer database.

Introduction

Campylobacteri jejuni is a gram-negative bacteria containing a flagellum which causes it to move in a corkscrew like pattern. It is one of the most common causes of food poisoning. It is microaerophilic, meaning the organism needs oxygen because it cannot ferment or respire anaerobically. *C. jejuni* is sensitive to its environment because too much oxygen can kill it.

In the United States, studies have shown that it is the leading cause of bacterial diarrheal illness. The bacteria contaminates raw chicken, non-chlorinated water, and raw milk. This organism is repeatedly found in the intestinal tracts of cats, dogs, cattle, poultry, rodents, and some humans. Even the smallest doses of *C. jejuni* can cause sickness two to five days after ingesting contaminated food or water. The illness, called campylobacteriosis, lasts from two to ten days. However, it is possible to die from the infection, causing an estimate of 124 deaths per year. Symptoms include diarrhea, fever, abdominal pain, nausea, headaches, or muscle pain. Complications can include meningitis, urinary tract infections; or rarely, short-term arthritis and Guillain-Barre syndrome (paralysis). Although anyone can get it, children under five, young adults, and elders are more prone to the infection.

To prevent the disease, people need to properly cook food as well as separate raw meat, poultry, and seafood from other foods to prevent cross- contamination. With the right control, *Campylobacter jejuni* can be minimized and prevented.



Scanning electron micrograph of *Campylobacter jejuni*. Credit: De Wood, Pooley, USDA's Agricultural Research Service. *medicalxpress.com*

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:

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?

Results

CJE0498		
Basic Information	**DNA Coordinates**	complement(453301..454398)
	DNA Length	1098 bp
	Amino Acid Sequence Length	365 aa
Sequence Based Similarity	**COGs**	COG0840:Methyl-accepting chemotaxis protein [Cell motility, Signal transduction mechanisms]
	TCOFFEE	not well conserved for most of the protein
	WEBLOGO	not well conserved in beginning and middle, but well conserved 1131-1159
Cellular Localization	**Transmembrane Helices**	none
	Signal Peptide	none
	PSORTB Final Prediction	unknown
Structure Based Evidence	**TIGRFAMs**	no significant hits
	PFAMs	Methyl-accepting chemotaxis protein (MCPSignal)
	HMM Logo Key Residues	G64, G67, G69, F70
	PDB	no significant hits

Campylobacter jejuni CJE0498:

COG and PFAM data predict CJE0498 to be a methyl-accepting chemotaxis protein. This type of protein can sense molecular changes in the extracellular matrix, allowing the bacteria to adjust its movements accordingly. Although Cellular Localization data was inconclusive, it can be inferred that this protein resides in the membrane based on its proposed function.

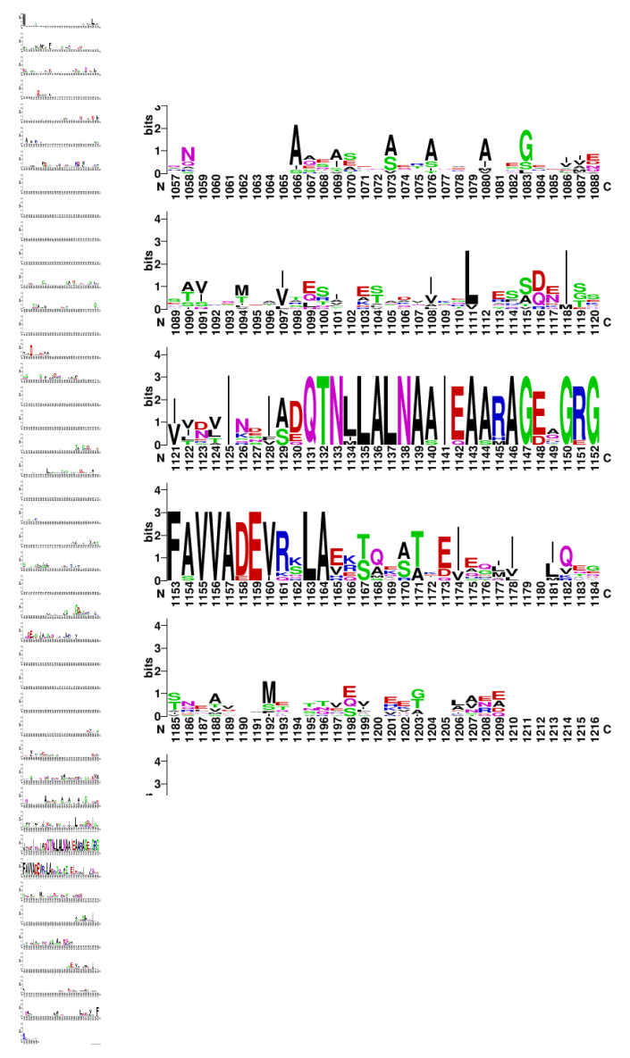


Figure I – WebLogo Data for *Campylobacter jejuni CJE0498*
The first image to the right shows the entire WebLogo and the protein's minimal conservation. The image on the far right shows the section that is conserved from amino acid 1131-1159.

CJE0699		
Basic Information	**DNA Coordinates**	629574..630395
	DNA Length	822 bp
	Amino Acid Sequence Length	273 aa
Sequence Based Similarity	**COGs**	COG0760: SurA Parvulin-like peptidyl-prolyl isomerase [Posttranslational modification, protein turnover, chaperones]
	TCOFFEE	Not well conserved in the beginning or end, more conserved in the middle
	WEBLOGO	Not well conserved in the beginning or end, more conserved in the middle
Cellular Localization	**Transmembrane Helices**	none
	Signal Peptide	yes: cleavage site between 21 & 22
	PSORTB Final Prediction	Outer membrane
Structure Based Evidence	**TIGRFAMs**	No significant TIGRFAMs
	PFAMs	Rotamase_3
	HMM Logo Key Residues	I21,F53, G71, S61, D63, G70, G106, H108
	PDB	3RFW: virulence factor PEB4 & periplasmic protein Cj1289: two structurally-related SurA-like chaperones in *C. jejuni*

Campylobacter jejuni CJE0699:

The data supports CJE0699 to be a peptidyl-prolyl isomerase which is an enzyme involved in the final folding and modification of a protein. CJE0699 converts the cis and trans isomers of peptide bonds with the amino acid proline. This reaction occurs outside the cell which is why the enzyme resides in the outer membrane and contains a signal peptide to be released when needed.

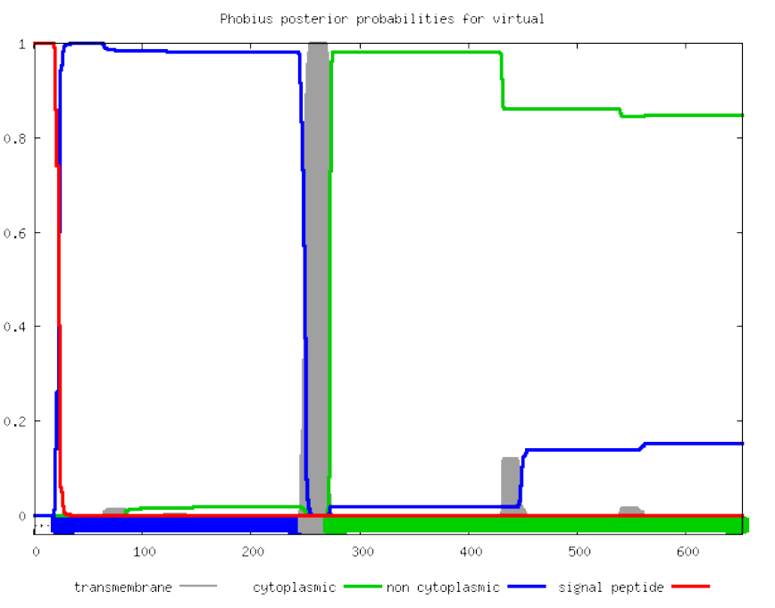


Figure II – Phobius Data for *Campylobacter jejuni* CJE0699.
The image to the right shows the presence of one transmembrane helix as well as a signal peptide indicated that the protein is predicted to be in the outer membrane of the bacteria and then released from the cell.

CJE0400		
Basic Information	**DNA Coordinates**	358928...359236
	DNA Length	309 bp
	Amino Acid Sequence Length	102 aa
Sequence Based Similarity	**COGs**	COG1886: FliN - Flagellar motor switch/type III secretory pathway protein FliN [Cell motility, Intracellular trafficking, secretion, and vesicular transport]
	TCOFFEE	Well conserved throughout
	WEBLOGO	Well conserved throughout
Cellular Localization	**Transmembrane Helices**	none
	Signal Peptide	none
	PSORTB Final Prediction	Cytoplasmic
Structure Based Evidence	**TIGRFAMs**	fliN: flagellar motor switch protein FliN
	PFAMs	Type III Flagellar Switch Regulator (C-Ring) FliN C-term
	HMM Logo Key Residues	G30
	PDB	FliM (SPOA):FliN fusion protein

Campylobacter jejuni CJE0400:

Data supports CJE0400 to be a flagellar motor switch protein found in the cytoplasm. It regulates the direction of flagellar motion, controlling locomotion in the bacteria. Due to the presence of many orthologs, it can be inferred that this is a common protein in many different species of bacteria containing a flagellum.

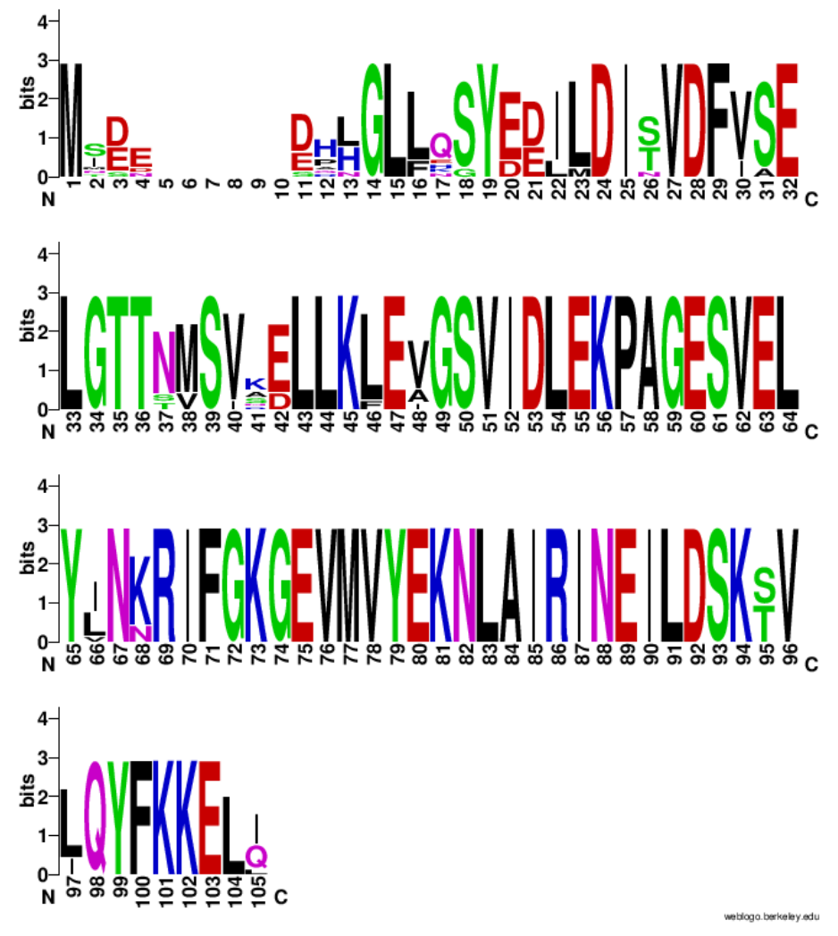


Figure III – WebLogo data for *Campylobacter jejuni* CJE0400.
The image to the right shows how well conserved the protein is with regards to amino acid sequence.

CJE0415		
Basic Information	**DNA Coordinates**	complement(370282..373404)
	DNA Length	3123 bp
	Amino Acid Sequence Length	1040 aa
Sequence Based Similarity	**COGs**	no significant hits
	TCOFFEE	Well conserved throughout
	WEBLOGO	Well conserved throughout
Cellular Localization	**Transmembrane Helices**	12
	Signal Peptide	none
	PSORTB Final Prediction	Cytoplasmic membrane
Structure Based Evidence	**TIGRFAMs**	heavy metal efflux pump, CzcA family
	PFAMs	No significant hit
	HMM Logo Key Residues	NA
	PDB	No significant hit

Campylobacter jejuni CJE0415:

TIGRFAM data supports CJE0415 to be a heavy metal efflux pump found on the cytoplasmic side of the membrane. The pump removes toxins from inside the bacterium *C. jejuni* allowing it to possibly be heavy metal resistant. TMHMM data shows it has 12 transmembrane helices.

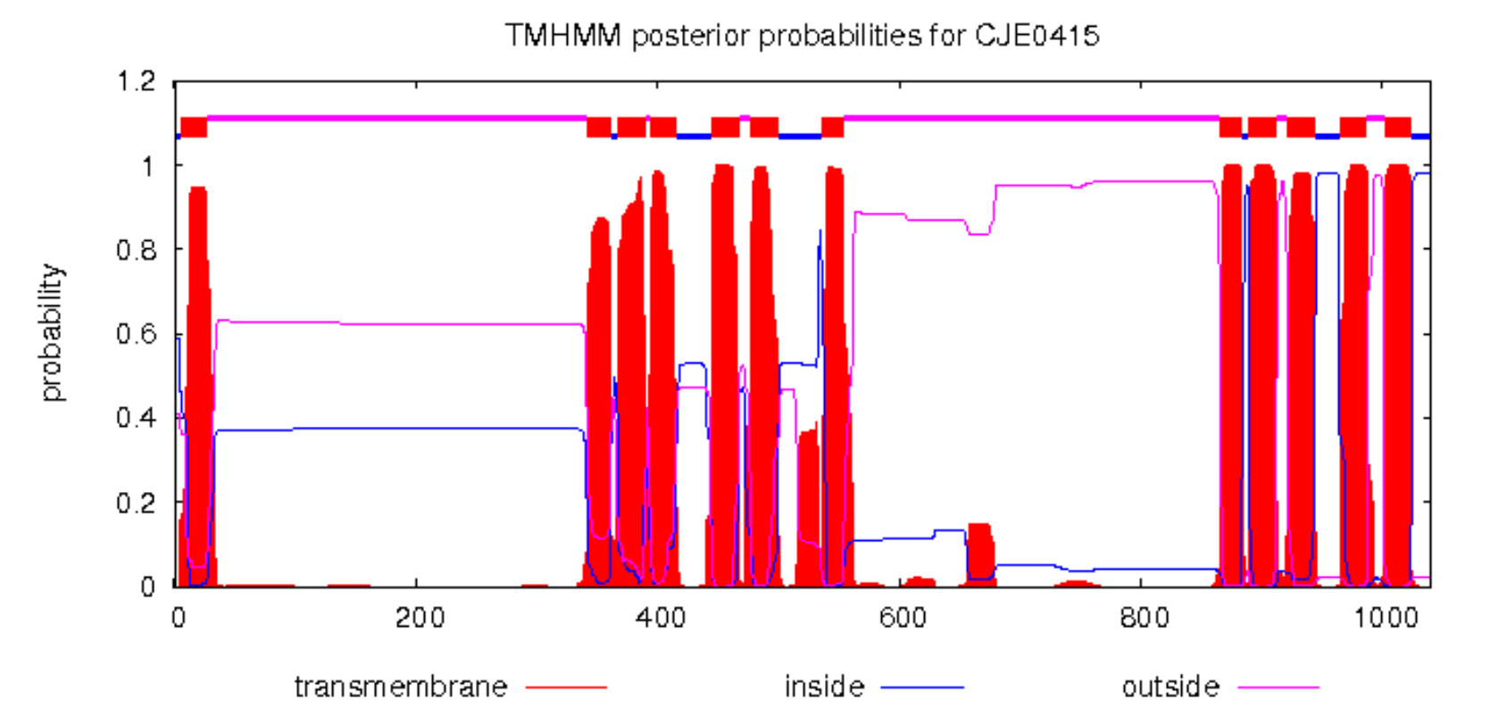


Figure IV – Cellular Localization from the TMHMM Database for *Campylobacter jejuni* CJE0415.
The image above shows the presence of twelve transmembrane helices.

CJE0820		
Basic Information	**DNA Coordinates**	complement(754967..755716)
	DNA Length	750 bp
	Amino Acid Sequence Length	249 aa
Sequence Based Similarity	**COGs**	COG1344: FlgL (flagellum)
	TCOFFEE	Conserved in the middle
	WEBLOGO	Sporadically conserved throughout
Cellular Localization	**Transmembrane Helices**	none
	Signal Peptide	none
	PSORTB Final Prediction	Periplasmic
Structure Based Evidence	**TIGRFAMs**	no significant hits
	PFAMs	Bacterial flagellin N-terminal helical region
	HMM Logo Key Residues	G31
	PDB	3K8V: Crysatl structure of a bacterial cell-surface flagellin N20C20

Campylobacter jejuni CJE0820:

COG, PDB, and PFAM data supports CJE0820 to play a role in the making of the flagellum of the bacterium. It is predicted to be located in the periplasm which is between the cell membrane and the cell wall. The protein builds the flagellum's "roots" in the periplasm from which it can grow from.

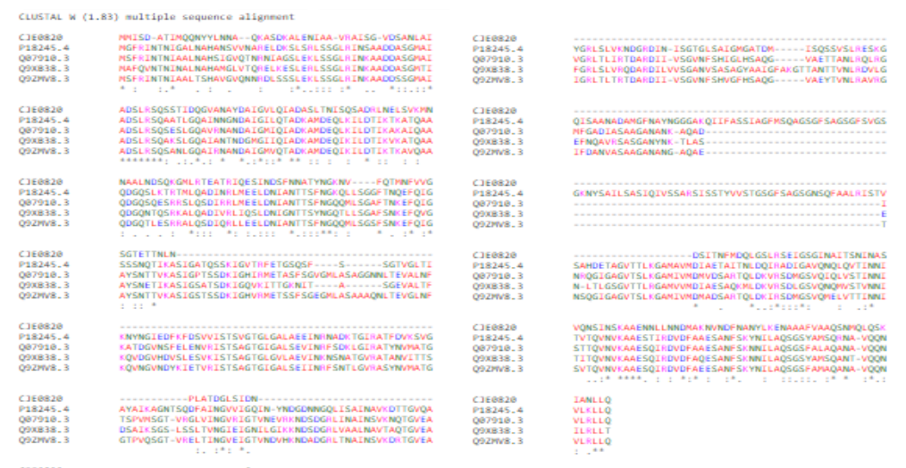


Figure V – TCOFFEE data for *Campylobacter jejuni* CJE0820.
The image shows the amino acid sequence to be mostly well conserved among other species of bacteria predicted to have this protein .

CJE0188		
Basic Information	**DNA Coordinates**	188197..189582
	DNA Length	1386 bp
	Amino Acid Sequence Length	461 aa
Sequence Based Similarity	**COGs**	no significant hits
	TCOFFEE	Well conserved in the middle and the end
	WEBLOGO	Well conserved in the middle and the end
Cellular Localization	**Transmembrane Helices**	none
	Signal Peptide	none
	PSORTB Final Prediction	Unknown
Structure Based Evidence	**TIGRFAMs**	FliI_clade1: flagellar protein export ATPas
	PFAMs	ATP-synt_ab: ATP synthase alpha/beta family, nucleotide-binding domain
	HMM Logo Key Residues	P204
	PDB	5B0O: Entity 1 containing Chain A, B, C, D Structure of the FliH-FliI complex

Campylobacter jejuni CJE0188:

TIGRFAM, PFAM, and PDB data predict CJE0188 to an ATP synthase protein involved in transporting materials used in flagellum assembly out of the cell. Although Cellular Localization data was inconclusive, one might predict this protein to be in the membrane based on its proposed function.

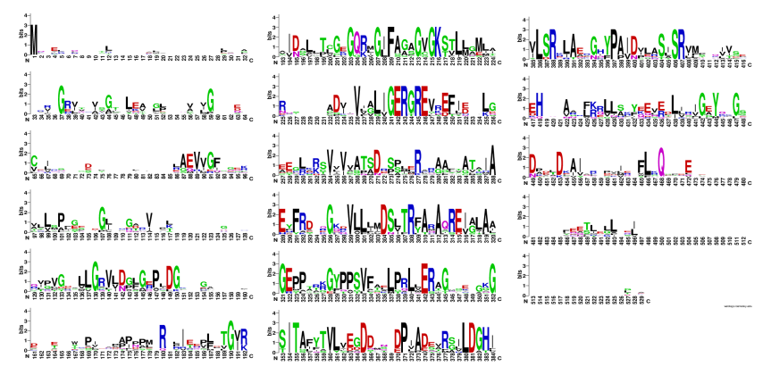


Figure VI – WebLogo Data for *Campylobacter jejuni* CJE0188 showing conservation in the middle of the protein.

Conclusion

The Geni-Act proposed gene product did not differ significantly from the proposed gene annotation for each of the genes and as such, the genes appear to be correctly annotated by the computer database.

References

Center for Food Safety and Applied Nutrition. (n.d.). Bad Bug Book - BBB - Campylobacter jejuni. Retrieved March 07, 2017, from https://www.fda.gov/food/foodborneillnesscontaminants/causesofillnessbadbugbook/ucm070024.htm

Acknowledgments

Supported by NSF ITEST Strategies Award Number 1311902

Gene Locus	Proposed Annotation
0498	Methyl-accepting chemotaxis protein
0699	Peptidyl-prolyl isomerase
0400	fliN: flagellar motor switch protein
0415	Heavy metal efflux pump
0820	Bacterial flagellin N-terminal helical region
0188	Flagellar protein export ATPas