Annotation of *Helicobacter pylori* at Locus Tags HP0169, HP0071, HP0913, and HP0887

Alyson Acquard, Daniel Gersitz, Lilian Koch, Bailey Stoldt, and Laura O'Donnell
Attica High School- Attica, 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

University at Buffalo

Abstract

A group of genes from the microorganism *Helicobacter pylori* 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

Helicobacter pylori is a gram-negative, spiral shaped bacteria that grows in the mucus layer of a human stomach. This bacteria can cause ulcers by breaking down the lining of the stomach and increases the risk of developing gastric cancer. Immune cells have a difficult time attacking *H. pylori* because they cannot reach the stomach lining. Also, *H. pylori* can interfere with immune cells and make them ineffective. To survive in the acidic environment of the stomach, *H. pylori* releases an enzyme that changes urea into ammonia. This conversion neutralizes the acidity and creates a more hospitable environment. The bacteria can also burrow into the mucus layer of the stomach and attach to the cells of the stomach lining.

An estimated ⅔ of the world's population has *H. pylori*. However, most people infected do not develop any symptoms. *H. pylori* spreads through contaminated food and water and direct mouth to mouth contact. Children living in poverty with unsanitary or crowded conditions are most likely to get infected. Ulcers can cause a burning pain in the stomach which worsens when the stomach is empty. Ulcers can also cause bloating, weight loss, and nausea. *H. pylori* can be detected through a breath test, blood test, or a stool examination. It can be treated with antibiotics or medications that reduce stomach acidity.

Colored scanning electron micrograph of *H pylori* on surface of gastric cells
web.stanford.edu

Methods

Modules of the GENI-ACT (http://www.geni-act.org/) were used to complete *Helicobacter pylori* 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- Structure-Based Evidence	TIGRfam, Pfam, PDB	Are there functional domains in my protein?
Module 4- Cellular Localization Data	Gram Stain, TMHMM, SignalP, PSORT, Phobius	Is my protein in the cytoplasm, secreted or embedded in the membrane?

Results

Gene Locus: HP0169		
Basic Information	DNA Coordinates	Complement (175453..176721)
	DNA Length	1269 nt
	Amino Acid Sequence Length	422 aa
Sequence-based Similarity	COGS	COG0826: PrtC- Collagenase-like protease
	T-Coffee	well conserved at beginning; not at end
	Web-Logo	Sporadic conservation throughout
Structure-based Evidence	TIGRFAMs	no significant hits
	PFAMs	pfam01136: Peptidase_U32
	HMM Logo Key Residues	E88,Y192
	PDB	no significant hits
Cellular Localization	Transmembrane Helices	none
	Signal Peptide	none
	PSORTb Final Prediction	cytoplasmic

Helicobacter pylori HP0169:

COG and PFAM data supports HP0169 to be a collagenase-like protease. It is an enzyme that breaks down proteins. Cellular localization data predicts that this protein functions in the cytoplasm.

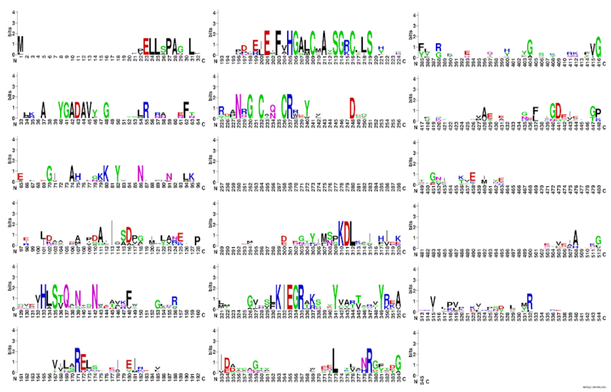

Figure 1 – WebLogo data for HP0169 shows sporadic areas of conservation of the protein amongst different species of bacteria.

Gene Locus: HP0071		
Basic Information	DNA Coordinates	Complement (74747..75334)
	DNA Length	588 nt
	Amino Acid Sequence Length	195aa
Sequence-based Similarity	COGS	no significant hits
	T-Coffee	no orthologs
	Web-Logo	no Web-Logo
Structure-based Evidence	TIGRFAMs	no significant hits
	PFAMs	AmiS/UreI family transporter
	HMM Logo Key Residues	P158, F67
	PDB	Urea Channel
Cellular Localization	Transmembrane Helices	6
	Signal Peptide	none
	PSORTb Final Prediction	Cytoplasmic membrane

Helicobacter pylori HP0071:

PFAM, and PDB data supports HP0071 to be a proton-gated urea channel that allows urea to be broken down to neutralize acid. Due to no orthologs present, it can be inferred that this protein is unique to *H. pylori* and is an adaptation to survive in the acidic stomach. Cellular localization data predicts that HP0071 is a cytoplasmic membrane protein.

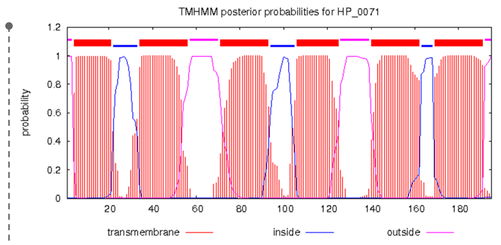

Figure 2 – TMHMM data for HP0071 The cellular localization data shows the presence of 6 transmembrane helices.

Gene Locus: HP0913		
Basic Information	DNA Coordinates	966746..968335
	DNA Length	1590 nt
	Amino Acid Sequence Length	529 aa
Sequence-based Similarity	COGS	No significant hits
	T-Coffee	No orthologs
	Web-Logo	No significant hits
Structure-based Evidence	TIGRFAMs	No significant hits
	PFAMs	pfam01856: HP_OMP Helicobacter outer membrane protein
	HMM Logo Key Residues	G5, R19, G70, F102, G109, G123, F158
	PDB	No significant hits
Cellular Localization	Transmembrane Helices	None
	Signal Peptide	Yes
	PSORTb Final Prediction	Outer membrane

Helicobacter pylori HP0913:

PFAM data suggests that HP0913 is an outer membrane protein. Little evidence exists as to the function of this membrane protein. Due to no orthologs given, it can be inferred that this protein may be unique to *H. pylori*. Cellular localization data predicts that HP0913 may contain a signal peptide cleaved from the outer membrane.

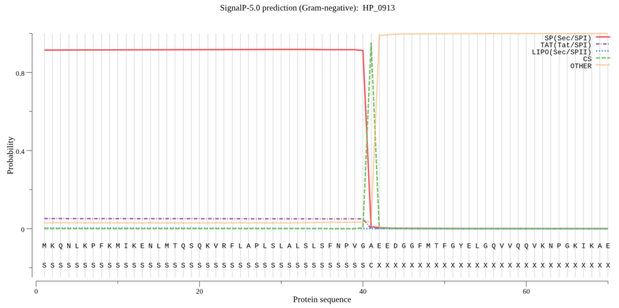

Figure 4 – SignalP data for *Helicobacter pylori* HP0913 showing the presence of a possible signal peptide with a likely cleavage site between position 41 and 42

Gene Locus: HP0887		
Basic Information	DNA Coordinates	938415..942287
	DNA Length	3873 nt
	Amino Acid Sequence Length	1290 aa
Sequence-based Similarity	COGS	COG4625: Uncharacterized conserved protein, contains a C-terminal beta-barrel porin domain
	T-Coffee	No orthologs
	Web-Logo	No Web-Logo
Structure-based Evidence	TIGRFAMs	TIGR01414: autotrans_barl:contain a conserved C-terminal domain that integrates into the outer membrane and enables the N-terminal region to be delivered across the membrane.
	PFAM	pfam03797: Autotransporter beta-domain
	HMM Logo Key Residues	W2, G31, G43, P138
	PDB	No significant hits
Cellular Localization	Transmembrane Helices	1
	Signal Peptide	Yes
	PSORTb Final Prediction	Outer Membrane/ Extracellular

Helicobacter pylori HP0087:

Lack of experimental evidence shows HP0087 to be a hypothetical protein. COG, TIGRFAM, and PFAM data indicate that at least this protein may reside in the outer membrane as an autotransporter. Cellular localization data predicts that HP0087 has 1 transmembrane helix as well as a signal peptide.

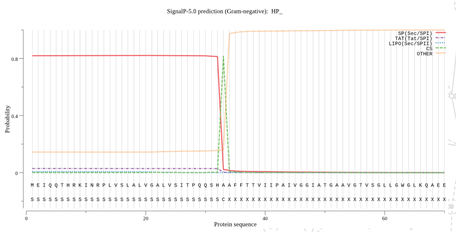

Figure 4 – SignalP data for *Helicobacter pylori* HP0087 showing the presence of a possible signal peptide with a likely cleavage site between position 33 and 34

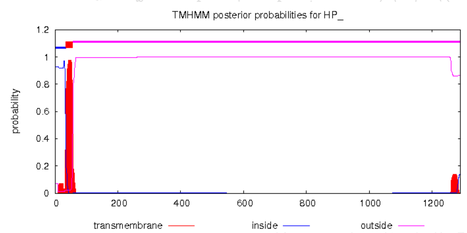

Figure 5 – TMHMM data for *Helicobacter pylori* HP0087. The cellular localization data shows the presence of 1 transmembrane helix

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. All proteins annotated seem to be involved with *H. pylori's* pathogenic nature and ability to survive in the stomach.

Gene Locus	Geni-Act Gene Products	Proposed Annotation
HP0169	PrtC- Collagenase-like protease	PrtC- Collagenase-like protease
HP0071	AmiS/UreI family transporter	AmiS/UreI family transporter
HP0913	Outer membrane protein	Outer membrane protein
HP0087	Hypothetical protein- possible autotransporter	Hypothetical protein- possible autotransporter

References

H. Pylori | Helicobacter Pylori Infections. (2019, April 08). Retrieved from https://medlineplus.gov/helicobacterpyloriinfections.html

Helicobacter pylori and Cancer. (n.d.). Retrieved from https://www.cancer.gov/about-cancer/causes-prevention/risk/infectious-agents/h-pylori-fact-sheet

Acknowledgments

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

www.buffalo.edu