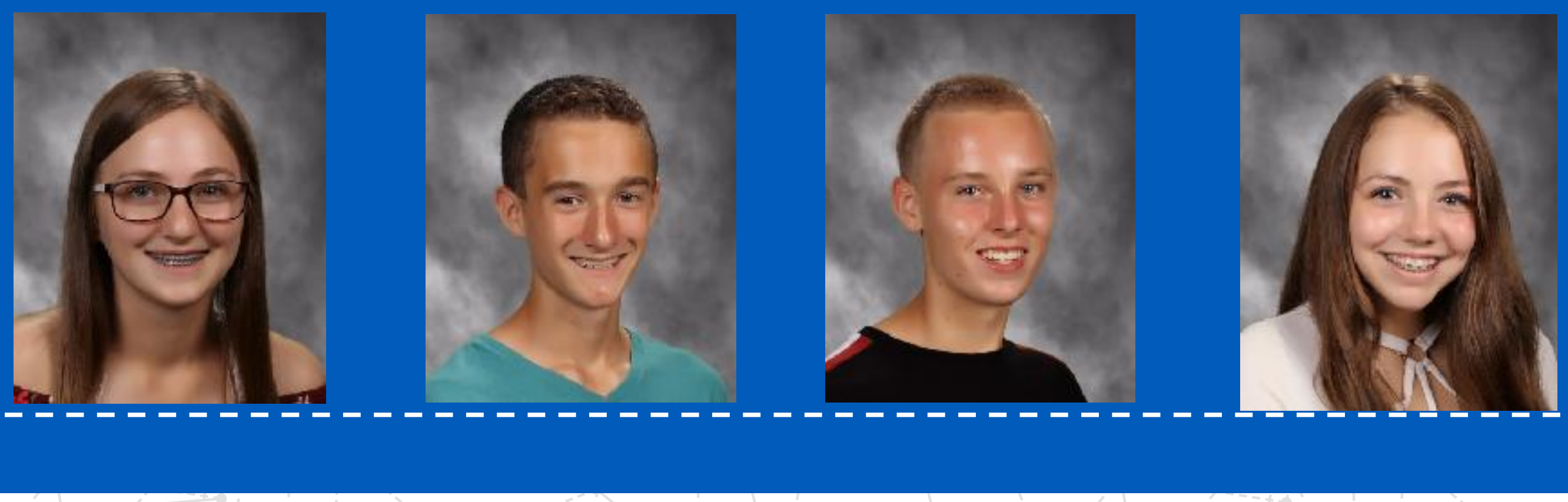
Annotation of *Helicobacter pylori* at Locus Tags HP0176, HP0071, HP0913, and HP0887

Gwenyth Cooley, Kyle George, Edward Janes, Reilly Quinn, 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, also known as *H. pylori*, is a gram-negative, spiral-shaped bacterium that are found in mucus layers in the lining of the stomach. *H. pylori* moves by the means of multiple flagella and use them to burrow into the mucus lining of the stomach.

Helicobacter pylori is commonly found in the majority of people. About two-thirds of the world's population has it in their body, but can remain dormant. If the bacteria do become active, they can cause sores, called ulcers, in the lining of your stomach and the duodenum. *H. pylori* causes more than 90% of duodenal ulcers and up to 80% of gastric ulcers which can sometimes lead to stomach cancer. The bacteria releases an enzyme called urease which reduces stomach acidity, allowing the bacteria to survive in the stomach.

There are many ways to test one for the presence of *Helicobacter pylori* including blood antibody tests, urea breath tests, stool antigen tests, and endoscopic biopsies. Although the bacteria can cause problems, there are antibiotics that can kill the bacteria and help the ulcers heal. The bacteria may be passed from person to person by direct contact of saliva, vomit, or fecal matter. It can also be spread by consuming contaminated food and water.

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: **HP0176**		
Basic Information	**DNA Coordinates**	182781..183704
	DNA Length	924 nt
	Amino Acid Sequence Length	307 aa
Sequence-based Similarity	**COGS**	COG0191:Fba-Fructose/tagatose bisphosphate aldolase [Carbohydrate transport and metabolism]
	T-Coffee	well conserved at the beginning not the end
	Web-Logo	fairly well conserved throughout
Structure-based Evidence	**TIGRFAMs**	TIGR01859 fruc_bis_ald_: class II fructose-1,6-bisphosphate aldolase
	PFAMs	pfam01116 F_bP_aldolase
	HMM Logo Key Residues	D84 H85 G177 H207
	PDB	class Ⅱ fructose-biphosphate aldolase
Cellular Localization	**Transmembrane Helices**	none
	Signal Peptide	none
	PSORTb Final Prediction	cytoplasmic

Helicobacter pylori HP0176:
COG, PFAM, TIGRFAM, and PDB data supports HP0176 to be a fructose-bisphosphate aldolase, an enzyme that breaks down fructose 1,6-bisphosphate into DHAP and glyceraldehyde 3 phosphate during glycolysis for energy metabolism. Cellular localization data predicts this protein to function in the cytoplasm.

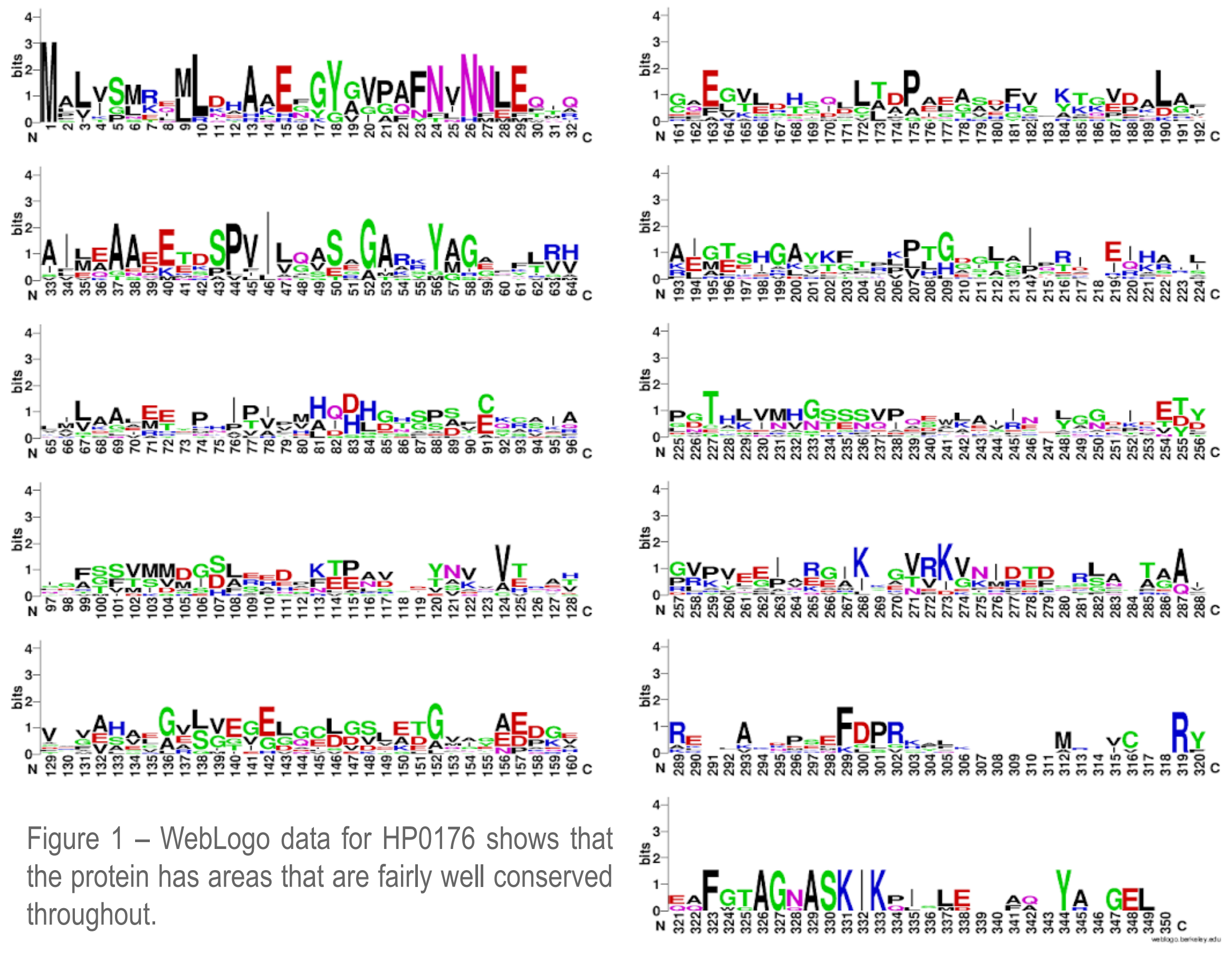


Figure 1 – WebLogo data for HP0176 shows that the protein has areas that are fairly well conserved throughout.

Gene Locus: **HP1166**		
Basic Information	**DNA Coordinates**	1230660..1232297
	DNA Length	1638 nt
	Amino Acid Sequence Length	545 aa
Sequence-based Similarity	**COGS**	No significant hits
	T-Coffee	No orthologs
	Web-Logo	No orthologs
Structure-based Evidence	**TIGRFAMs**	No significant hits
	PFAMs	PF00342: Phosphoglucose isomerase
	HMM Logo Key Residues	No significant residues
	PDB	pGI glucosephosphate isomerase
Cellular Localization	**Transmembrane Helices**	none
	Signal Peptide	none
	PSORTb Final Prediction	cytoplasmic

Helicobacter pylori HP1166:
PFAM and PDB data suggests that HP1166 is phosphoglucose isomerase protein which converts glucose 6-phosphate into fructose 6-phosphate during glycolysis for energy metabolism. Cellular localization data predicts that HP1166 functions in the cytoplasm.

Gene Locus: **HP1275**		
Basic Information	**DNA Coordinates**	1350206...1351585
	DNA Length	1380 nt
	Amino Acid Sequence Length	459 aa
Sequence-based Similarity	**COGS**	COG1109: ManB-Phosphomannomutase [Carbohydrate transport and metabolism]
	T-Coffee	Not well conserved in the beginning and end. Well conserved in the middle.
	Web-Logo	Mostly well-conserved throughout
Structure-based Evidence	**TIGRFAMs**	TIGR03990:Arch_GlmM; catalyzes the conversion of glucosamine-6-phosphate to glucosamine-1-phosphate.
	PFAMs	pfam02878:PGM_PMM_I; Phosphoglucomutase/ phosphomannomutase, alpha/beta/alpha domain I
	HMM Logo Key Residues	D48
	PDB	PMM/PGM with ribose 1-phosphate
Cellular Localization	**Transmembrane Helices**	none
	Signal Peptide	none
	PSORTb Final Prediction	cytoplasmic

Helicobacter pylori HP1275:
COG, TIGRFAM, PFAM, and PDB data supports HP1275 to be a phosphomannomutase, an enzyme that transports phosphate groups between sugar substrates, glucosamine-6 and glucosamine-1 in glycolysis for energy metabolism. Cellular localization data predicts that HP1275 functions in the cytoplasm.

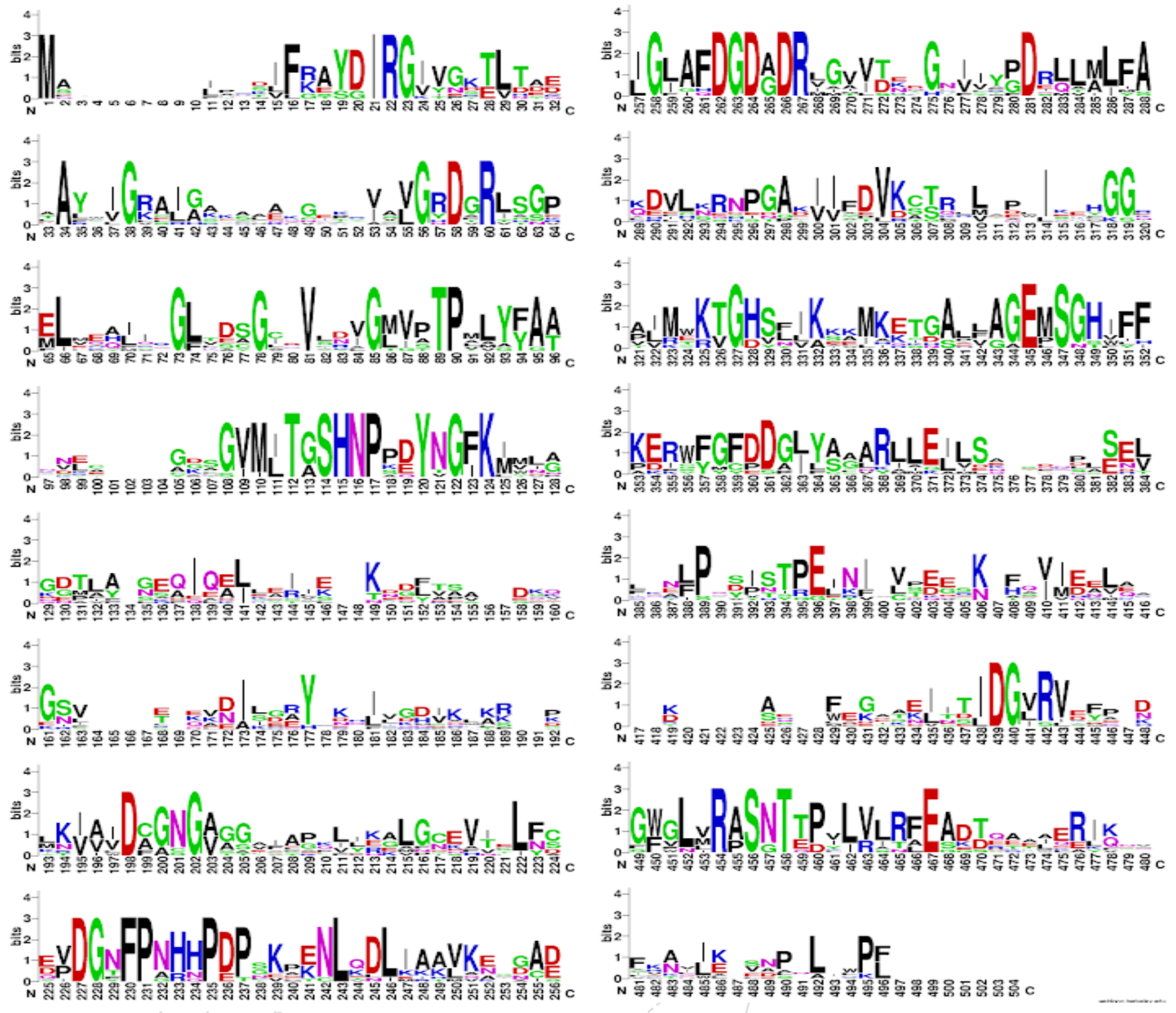


Figure 2 – WebLogo data for HP1275 shows that the protein has many well conserved areas throughout.

Gene Locus: **HP1358**		
Basic Information	**DNA Coordinates**	complement (1447156..1448028)
	DNA Length	873 nt
	Amino Acid Sequence Length	290aa
Sequence-based Similarity	**COGS**	COG0158: Fbp- Fructose-1,6 bisphosphatase [Carbohydrate transport and metabolism]
	T-Coffee	well conserved in the beginning, not as well conserved at the end
	Web-Logo	not well conserved
Structure-based Evidence	**TIGRFAMs**	no significant hits
	PFAMs	pfam00316: Fructose-1-6-bisphosphate, N-terminal domain
	HMM Logo Key Residues	G49, G58, P110, Y158
	PDB	sedoheptulose-1,7-bisphosphatase
Cellular Localization	**Transmembrane Helices**	none
	Signal Peptide	none
	PSORTb Final Prediction	Cytoplasmic

Helicobacter pylori HP1358:
COG, PFAM, and PDB data supports HP1358 to be a fructose-1,6-bisphosphatase, an enzyme that converts fructose-1,6-bisphosphate to fructose 6-phosphate in glycolysis for energy metabolism. Cellular localization data predicts that HP1358 functions in the cytoplasm.

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 energy metabolism in *H. pylori.*

Gene Locus	Geni-Act Gene Products	Proposed Annotation
HP0176	Fructose-bisphosphate aldolase	Fructose-bisphosphate aldolase
HP1166	Phosphoglucose isomerase	Phosphoglucose isomerase
HP1275	Phosphomannomutase	Phosphomannomutase
HP1358	Fructose-1,6-bisphosphatase	Fructose-1,6-bisphosphatase

References

H. pylori Bacteria Infection: Symptoms, Diagnosis, Treatment, Prevention. (n.d.). Retrieved from https://www.webmd.com/digestive-disorders/h-pylori-helicobacter-pylori#1

Helicobacter Pylori. (n.d.). Retrieved from https://www.cedars-sinai.org/health-library/diseases-and-conditions/h/helicobacter-pylori.html

Acknowledgments

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

www.buffalo.edu