Annotation of the *Helicobacter Pylori 83* Genome from DNA Coordinates 64777 to 73048

Gavin Bathgate, Amanda Broker and Lawrence Hohl

East Rochester High School, East Rochester 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

Our group annotated 3 genes from *Helicobacter pylori 83* genome using the modules given by GENI-ACT. Following the instructions given in our GENI-ACT manual, we went to many different sites that aided us in annotating our genes. These included BLAST, CDD, T-coffee, WebLogo, Gram Stain, TMHMM, SignalP, Psort, Phobius, IMG EDU, TIGRfam, Pfam, PDB, KEGG, MetaCyc and Rfam. We then recorded our findings in the online GENI-ACT lab notebook. Our results were established using all 2 students annotations. Finally, as you can see, we are presenting our results on this trifold poster. We added pictures and diagrams to help show what our genes do. In the end, most of the proposed annotations for the sequenced genomes were the same. The purpose of our research was to annotate specific genes then compare the results to those already given by GENI-ACT and see if the data proposed annotations were either correct or incorrect.

Introduction

The gram negative *Helicobacter pylori* 83 bacterium can alter the human regulatory mechanisms for gastric acid production. For many years, people believed that smoking, stress, spicy food, and other lifestyle caused stomach ulcers. It was later found that *Helicobacter pylori* 83 was ingested through unclean food. The pylorus is the sphincter muscles between the duodenum and the stomach. This research is very important because it will help us learn new ways to fight diseases and bacteria, such as *Helicobacter pylori* 83. *Helicobacter pylori* 83 can decrease acid secretion levels in the stomach. This can result in ulcers in the stomach, and even in the duodenum. Some signs of a *H. pylori* infection are an ache or burning pain in your abdomen, abdominal pain that's worse when your stomach is empty, nausea, loss of appetite, frequent burping, vomiting, unintentional weight loss, and bloating.

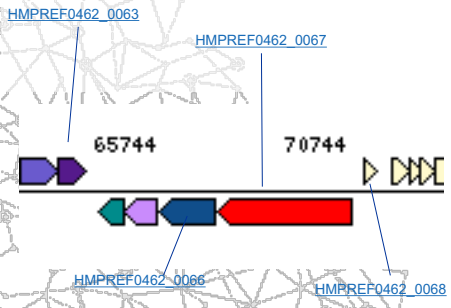

Figure 1. The locus tags and relative position of the genes under investigation in this research

Methods

Modules of the GENI-ACT (http://www.geni-act.org/) were used to complete *Helicobacter pylori 83* 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?
Module 6- Enzymatic Function	KEGG, MetaCyc, E.C. Number,	In what process does my protein take part?
Module 7- Gene Duplication/ Gene Degradation	Paralog, Pseudogene	Are there other forms of my gene in the bacterium? Is my gene functional?
Module 8- Evidence for Horizontal Gene Transfer	Phylogenetic Tree,	Has my gene co-evolved with other genes in the genome?
Module 9- RNA	RFAM	Does my gene encode a functional RNA?

Results

HMPREF0462_0063 :
The initial proposed product of this gene by GENI-ACT was a probable membrane protein TerC. This product proposal was supported by the top BLAST hits of its amino acid sequence as well as IMG EDU ortholog gene neighborhood, and COG.

HMPREF0462_0067 :
The initial proposed product of this gene by GENI-ACT was a 1-pyrroline-5-carboxylate dehydrogenase. This product proposal was supported by the top BLAST hits of its amino acid sequence as well as IMG EDU ortholog gene neighborhood, metaCyc and COG. Glutamic gamma-semialdehyde is the metabolic precursor for proline biosynthesis. The conversion from L-Glutamate, an ATP- and NADPH-dependent reaction, is catalyzed by the enzyme Delta-1-pyrroline-5-carboxylate synthetase.

HMPREF0462_0066 :
The initial proposed product of this gene by GENI-ACT was a Sodium/proline symporter. This product proposal was supported by the top BLAST hits of its amino acid sequence as well as IMG EDU gene ortholog neighborhood and COG all suggesting this is a proline symporter.

HMPREF0462_0068 :
The initial proposed product of this gene by GENI-ACT was a hypothetical protein. This product proposal was supported by the top BLAST hits of its amino acid sequence. For most modules there were no results

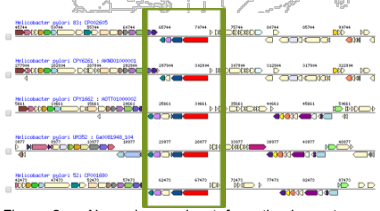

Figure 2 – Above is a snippet from the locus tag number HMPREF0462_0067 result from IMG EDU showing the ortholog gene neighborhood. Since there are very similar results we can conclude that there is no horizontal gene transfer

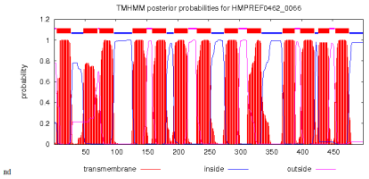

Figure 3 – Above is a snippet from the locus tag number HMPREF0462_0066 result from TMHMM showing posterior probabilities. This snippet shows a possible N-term signal sequence.

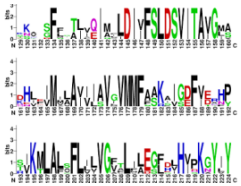

Figure 4 – Above is a snippet from the locus tag number HMPREF0462_0063 cropped result from WebLogo. The whole alignment shows that the proteins in the alignment are conserved throughout.

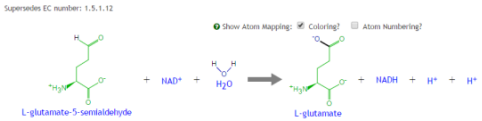

Figure 5 – Above is a snippet from the locus tag number HMPREF0462_0067 result from MetaCyc that this gene converts L-Glutamate-5-semialdehyde, an ATP, to L-Glutamate which is not an essential but is the most common excitatory neurotransmitter in the central nervous system and it is then transported.

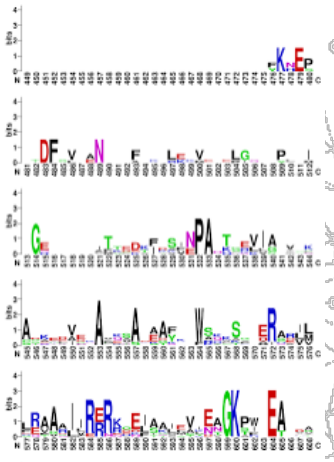

Figure 6 – A snippet from the locus tag number HMPREF0462_0067 cropped result from WebLogo. The whole alignment shows that the proteins in the alignment are not well conserved throughout.

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.

Gene Locus	Geni-Act Gene Products	Proposed Annotation
04170	Sodium/Proline Symporter	Sodium/Proline Symporter
04180	Histidine Ammonia-Lyase	Histidine Ammonia-Lyase
04190	Uncharacterized Anaerobic Dehydrogenase	F4/4S4 Domain / Alpha Subunit / Formate Dehydrogenase
04200	Uncharacterized Anaerobic Dehydrogenase	Molybdopterin Dinucleotide Binding Domain / Alpha Subunit / Formate Dehydrogenase
04210	Beta Subunit / Formate Dehydrogenase	Beta Subunit / Formate Dehydrogenase
04220	Formate-Dependent Nitrite Reductase Membrane Component	Formate-Dependent Nitrite Reductase
04230	Selenophosphate Synthase	Selenide,Water Dikinase

References

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

Acknowledgments

Supported by an NIH Science Education Partnership (SEPA) Award - R25ODO10536, Stephen Koury, Ph.D., Rama Dey-Rao, Ph.D.
www.buffalo.edu