Annotation of the *Wolbachia pipietis* Genome at Locus Tags WP0020 and WP0025-WP0027

Kourtney Ames, Reece Bates, Gabby Carris, Abigail Ferreira, Teddi Hayes, Joeseph Leamer, Haley Silka and Lennart Liffner
Sherman Central School, 127 Park St. Sherman NY 14781 and The Western New York Genetics in Research and Health Care Partnership

SEPA SCIENCE EDUCATION PARTNERSHIP AWARD
Supported by the National Institutes of Health

Abstract

Four genes from the microorganism *Wolbachia pipietis* (WP0020, WP0025-WP0027) 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, cellular localization data, potential alternative open reading frames, enzymatic function, presence or absence of gene duplication and degradation, the possibility of horizontal gene transfer, and the production of an RNA product. 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.

Introduction

Wolbachia pipietis are a group of gram negative bacteria, commonly living inside the cells of 17-67% of arthropods and various nematodes. This type of bacteria is transmitted through the reproduction of its host. It is passed on through the cytoplasm of the eggs.

Wolbachia manipulates it's host into parthenogenesis induction, feminization, male killing, and cytoplasmic incompatibility. All of these manipulations promote and increase in the female population. this is important ant because the *Wolbachia* bacteria is passed on only through females. Males that are infected with *Wolbachia* cannot pass down *Wolbachia* unless the female is already infected. Yet non-infected males can reproduce with both infected and non-infected females.

Although many host insects of *Wolbachia* are able to pass on many different human diseases such as West Nile virus and Lyme disease, *Wolbachia* is not infectious to humans. According to Prevot, *Wolbachia* has been researched through (PCR), and biochemical technology.

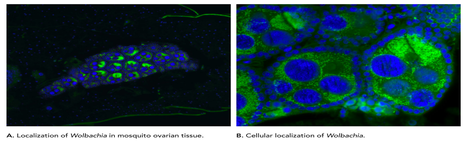

Figure 1. Wolbachia is shown in green in the Ovaries of the *Aedes aegypti* Mosquito.
https://www.idtdna.com/pages/education/decoded/article/dengue-zika-transmission-slowed-by-em-Wolbachia-em-bacterium

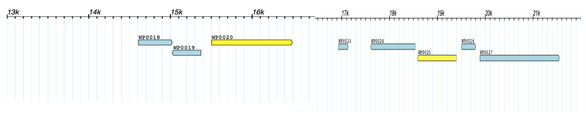

Figure 2. The locus tags and relative position of the genes under investigation in this research

Methods

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

Modules	Activities	Questions Investigated
Basic Information	DNA Coordinates and Sequence, Protein Sequence	What is the sequence of the gene and protein? Where is it located in the genome?
Sequence-Based Similarity	Blast, CDD, T-Coffee, WebLogo	How similar is the protein under investigation to other proteins in GenBank?
Structure-Based Similarity	TIGRfam, Pfam, PDB	What functional domains are present in the protein under investigation?
Cellular Localization	Gram Stain, TMHMM, SignalP, LipoP, Psortb, Phobius	Is the protein under investigation located in the cytoplasm, secreted, in the periplasm or embedded in the cell membrane or cell wall?
Enzymatic Function	KEGG, MetaCyc, E.C. Number	In what process or structure is the protein under investigation involved?
Duplication and Degradation	Paralog, Pseudogene	Are there other forms of the protein under investigation in the same genome? Is it functional?
Horizontal Gene Transfer	Phylogenetic Tree, Gene Neighborhood, GC Content	Has the protein under investigation co-evolved with the rest of the genome or has it been obtained in a different way?
RNA family	Rfam	Does the gene under investigation encode a functional RNA?
Final Annotation	Evaluate data from all modules	Has the gene been correctly called by the pipeline annotation?

Results

WP0020:
The computer pipeline proposed product of this gene was a ribonucleoside reductase. The enzyme is responsible factor in the conversion of ribonucleotides to deoxyribonucleotides which builds DNA. This was supported by the top BLAST hit which came back with the same product name for the top hits as the product proposed by the computer pipeline. The ribonucleoside reductase is normally based in the cytoplasm. The TMHMM search showed that the enzyme has one sign of a transmembrane helix. Yet the SignalP, and LipoP searches lead to the conclusion that it is indeed found in the cytoplasm. This all proves that the acceptable name for this enzyme is a ribonucleoside reductase.

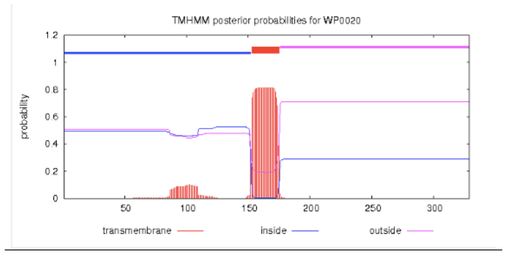

Figure 3. This is the TMHMM search which shows that there is one transmembrane helix indicated.

WP0025:
The computer pipeline proposed product of this gene was a aminomethyltransferase family protein. This was supported by the top BLAST hits for the amino acid sequences, the presence of well-curated functional domains within the amino acid sequences, the cellular location of the amino acid sequences, and the enzymatic function of the amino acid sequences. In the WebLogo the n terminus but there are regions of conservation scattered throughout the alignment. You would expect to find this protein in the cytoplasm. This gene was classified as an enzyme according to the data found. The acceptable name for this is Aminomethyltransferase family protein.

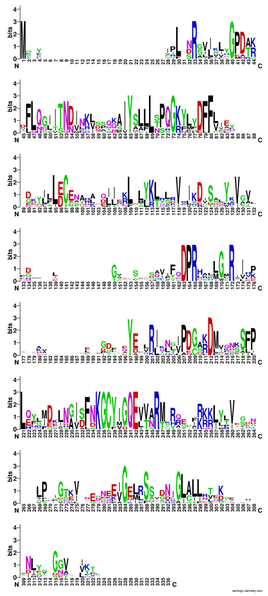

Figure 4. The WebLogo of the T-Coffee alignment from WP0025. Note scattered areas of conservation throughout the alignment

WP0026:
The computer pipeline depicted that my gene organism was *Wolbachia* Endosymbiont of *Drosophila melanogaster*. Typically these organisms are transmitted maternally from mother to daughter through the egg, although these bacteria can affect their hosts reproductive capabilities in order to enhance transmission. The blasts show that 10 kDa Chaperonin was the top hit. It's molecular function is to bind ATP.
The top hit COG name was Co-chaperonin GroES (HSP10). The WebLogo showed that from amino acid 1-75 there are a ton of similarities, but not many after except for 116-123 and 371-380. With all the data completed and collected it is clear that the gene WP0026 is named 10 kDa Chaperonin and the organism is *Wolbachia*.

RecName: Full=10 kDa chaperonin; AltName: Full=GroES protein; AltName: Full=Protein Cpn10
Sequence ID: C0R2V4.1 Length: 96 Number of Matches: 1

Range 1: 1 to 96 GenPept Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps
172 bits(437)	7e-57	Compositional matrix adjust.	82/96(85%)	92/96(95%)	0/96(0%)

```
Query  1   MSSVNLNVLDDNVLIKPITEEKQGGIVLPSSAEKKPTKGEVIAIGSGSRNSSGERIALTV  60
           MSS++LNVLDD+VLIKPI+EEKQGGIVLPSSAEKKPTKGEVIAIG GSRNSSGER+ LTV
Sbjct  1   MSSISLNVLDDSVLIKPISEEKQGGIVLPSSAEKKPTKGEVIAIGEGSRNSSGERVTLTV  60

Query  61  KTGDKVFYRQWAGTEVEHDNEKYVVMKESDLLAVIK  96
           K GDKVFYRQWAGTE+EH++EK +VMKESD+LAVIK
Sbjct  61  KAGDKVFYRQWAGTEIEHNDEKLIVMKESDILAVIK  96
```

Figure 5. WP0026 BLAST results revealing identities, positives, and gaps in the amino acid sequence

WP0027:
The computer pipeline proposed product of this gene was a chaperonin. This was supported by the top BLAST hits which came back with the same product name for the top hits as the product proposed by the computer pipeline. These results would lead to the conclusion that the given protein will look the same as *Wolbachia*.

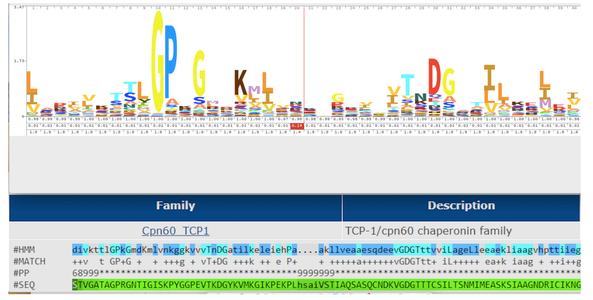

Figure 6. At the beginning several key functional residues are found with several scattered across the length of the protein.

Conclusions

The final results showed that the proposed annotations ended up being the same as the proposed gene products from GENI-ACT. There were no significant differences in any of the genes noted so our results were successful.

Locus Tag	Pipeline Annotation Product Name	Proposed Annotation	Changes Proposed?
WP0020	Ribonucleoside Reductase	Ribonucleoside Reductase	No
WP0025	Aminomethyl Transferase Family Protein	Aminomethyl Transferase Family Protein	No
WP0026	10 KDA Chaperonin	10 KDA Chaperonin	No
WP0027	60 KDA Chaperonin	60 KDA Chaperonin	No

References

Prevot, Karine, "Wolbachia". Embryo Project Encyclopedia (2015-01-). ISSN: 1940-5030 http://embryo.asu.edu/handle/10776/8285. 29

University at Buffalo The State University of New York
Jacobs School of Medicine and Biomedical Sciences
buffalo.edu

Acknowledgments

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