Annotation of the *Kytococcus sedentarius* Genome from Locus Tags Ksed_09640 to Ksed_09670

Lydia Borrelli, Dalton Cowell, Elijah Flansburg, Calvin Herrick, and Michael Chiulli
Le Roy Junior-Senior High School and the Western New York Genetics in Research Partnership

University at Buffalo

Abstract

A group of consecutive four genes from the microorganism Kytococcus sedentarius (Ksed_09640 – Ksed_09670) were annotated using the collaborative genome annotation website GENI-ACT. The Genbank proposed gene product name for each gene was assessed in terms of general genome information (DNA Coordinates and Sequence, protein Sequence), Sequence-Based Similarity Data (Blast, CDD, T-Coffee, Weblogo), Cellular Localization Data (Gram Stain, TMHMM, SignalP, PSORT, Phobius), Alternate Open Reading Frame (IMG Sequence Viewer), and Structure-Based Evidence (TIGRfam, Pfam, PDB). After the genes were assessed each 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 GENI-ACT nr database.

Introduction

Kytococcus sedentarius is a non-motile, aerobic, non-encapsulated, and gram positive bacterium. The bacterium was originally isolated in 1944 from a playground slide that was submerged in water, and it grows best in concentrations of sodium chloride of less than 10% (Sims et al., 2009). *Kytococcus sedentarius* has been used to produce antibiotics monensin A and monensin B (Sims et al., 2009). It has been known to cause pitted keratolysis, valve endocarditis, as well as hemorrhagic pneumonia (Sims et al., 2009). The bacterium is also of interest because it is a member of the *Dermacoccaceae* family that is within the suborder *Micrococcineae* which still has not been thoroughly studied for utilizing bioinformatics.

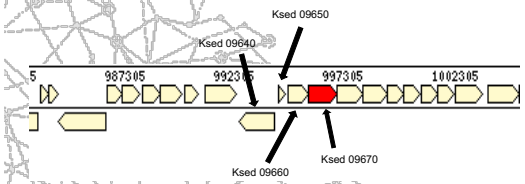

Figure 1.- Gene neighborhood of *Kytococcus sedentarius* for the four genes (Ksed_09640, Ksed_09650, Ksed_09660, Ksed_09670) studied.

Methods

Modules of the GENI-ACT (http://www.geni-act.org/) were used to complete *Kytococcus sedentarius* 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?

Results

Ksed_09640: The initial proposed product of this gene by GENI-ACT was catalase. This gene product proposal was supported by the top BLAST hits for the amino acid sequence, information found through TIGRFAM, Pfam, and PDB. As such, the proposed annotation is a catalase. Catalase is an enzyme commonly found in living organisms exposed to oxygen such as bacteria, plants, and animals. Most importantly, it catalyzes the decomposition of hydrogen peroxide to water and oxygen as well as protects cells from oxidative damage.

Ksed_09650: The initial proposed product of these genes by GENI-ACT was *trp operon repressor protein*. This gene product proposal was supported by the top BLAST hits for the amino acid sequences, the presence of well-curated functional domains within the amino acid sequences, and the cellular location of the amino acid sequences. TMHMM analysis indicated that no transmembrane helices were present, which makes sense because operon repressor proteins must be located in the cytoplasm to bind with DNA at a specific site to block transcription.

Ksed_09660: The initial proposed product of these genes by GENI-ACT was an ATP phosphoribosyltransferase which Catalyzes the condensation of ATP and 5-phosphoribose 1-diphosphate to form N'-(5'-phosphoribosyl)-ATP (PR-ATP).This was supported in both BLAST and TIGRFAM. There were no transmembrane helices according to the TMHMM data. ATP phosphoribosyltransferase is located in the cytoplasm. All data collected supports the discovery that Ksed_09660 is in fact ATP phosphoribosyltransferase.

Ksed_09670: The initial proposed product of these genes by GENI-ACT was a histidinol dehydrogenase. Top BLAST program hits for the amino acid sequence also support it to be histidinol dehydrogenase, as well as the existence of preserved domains in the sequence, and TMHMM localization data that determined there to be no transmembrane helices. No conflicting results to the GENI-ACT proposal were found, so the proposed annotation of Ksed_09670 is histidinol dehydrogenase.

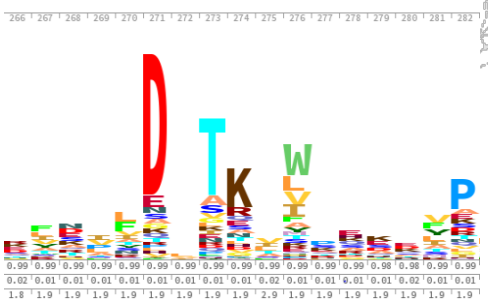

Figure 2 – A portion of the HMM logo for Ksed_09640. On the HMM logo, the larger letters represent more conserved amino acids. Colors correspond to different types of amino acids such as neutral and acidic. Letters are sorted in descending order depending on their probability of occurring at a given position in a sequence that contains the domain.

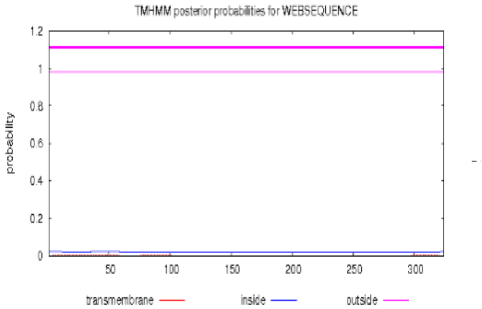

Figure 3 - The TMHMM graph for Ksed_09650 showing the absence of a transmembrane helix, which supports the proposal that the protein is a *trp* operon repressor protein. This is supported because an operon repressor must be within the cytoplasm in order for it to bind to the gene's operator region and block transcription.

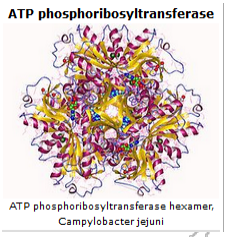

Figure 4 – Structure of ATP phosphoribosyltransferase. This enzyme catalyses the first step in the biosynthesis of histidine in bacteria, fungi and plants.

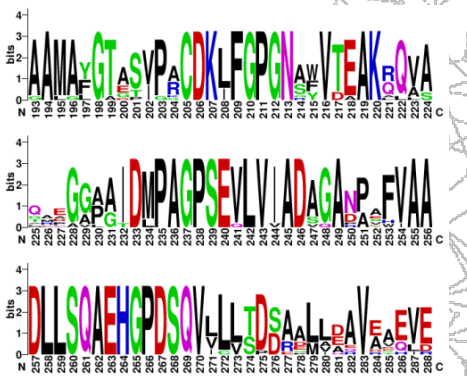

Figure 5 – Portion of the WebLogo for Ksed_09670 showing the higher degrees of conserved amino acids with larger letters, and amino acids that are less conserved with smaller letters.

Conclusion

The GENI-ACT proposed annotation for our gene products did not differ throughout the research.

Gene Locus	GENI-ACT Product	Proposed Annotation
Ksed_09640	Catalase	Catalase
Ksed_09650	Trp operon repressor	Trp operon repressor
Ksed_09660	ATP phosphoribosyltransferase	ATP phosphoribosyltransferase
Ksed_09670	histidinol dehydrogenase	histidinol dehydrogenase

References

Sims et al. (2009). Complete genome sequence of *Kytococcus sedentarius* type strain (541T). Standards Genomic Sciences,12 - 20.
Study.com. Study.com, n.d. Web. 16 May 2017.

Acknowledgments

Supported by an NSF Innovative Technology Experiences for Students and Teachers (ITEST) Award - 1311902

www.buffalo.edu