Annotation of the *Kytococcus sedentarius* Genome from Gene Locus Tags Ksed_08420 to Ksed_08450

Hayley Babbitt, Margaux Carmel, Elijah Flansburg, and Michael Chiulli

LeRoy Jr/Sr High School and the Western New York Genetics in Research Partnership

Abstract

A group of three genes in the same neighborhood from the microorganism *Kytococcus sedentarius* (Ksed_08420 through Ksed_08450) were annotated using the genome annotation website GENI-ACT. The proposed gene product information for each gene was assessed in terms of the general genomic information, amino acid sequence-based similarity data with the activities BLAST, CDD, T-Coffee, and WebLogo), structure-based evidence from the amino acid sequence (TIGRfam and Pfam), and cellular localization data (TMHMM, SignalP, and Phobius). The 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 in ther database.

Introduction

Kytococcus sedentarius is gram positive bacterium, and is found mainly in tetrad formation. *Kytococcus sedentarius* is classified as a chemoheterotroph, because amino acids are required for growth. It was originally isolated from a microscope slide submerged in seawater in 1944.

Kytococcus sedentarius is an interesting microorganism for many reasons. *Kytococcus sedentarius* is a natural source of the antibiotics monensin A and monensin B (Sims et al., 2009), which are both used in animal feeds to promote growth in the consumers. This bacterium has been identified as the cause of a number of opportunistic infections including valve endocarditis, hemorrhagic pneumonia, and pitted keratolysis (Sims et al., 2009). Finally, the phylogeny of this microorganism is a source of interest, as it is a member of the family *Dermacoccaceae* within the actinobacterial suborder *Micrococcineae*, which has not been thoroughly studied using bioinformatics (Sims et al., 2009).

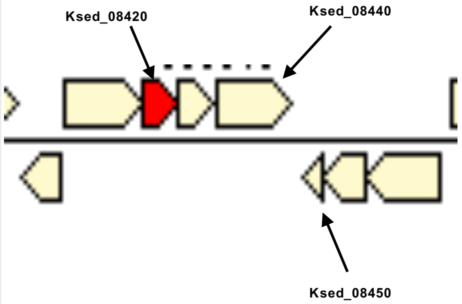

Figure I - Gene neighborhood of the three *Kytococcus sedentarius* genes; Ksed_08420, Ksed_08440, and Ksed_08450; annotated.

Methods and Materials

Four 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- 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

*Kytococcus sedentarius*08420:

The initial proposed product of this gene by GENI-ACT was a phosphoribosylaminoimidazole carboxylase. This gene product proposal was then supported by the cellular localization data presented by COG, PDB, and Pfam, along with the top BLAST hits for the amino acid sequence. This gene's function is most likely the aiding in nucleotide transport and metabolism, and the ability to convert AIR to CAIR directly and without ATP.

*Kytococcus sedentarius*08440:

The initial proposed product of this gene by GENI-ACT was an acyl-CoA dehydrogenase. This gene product proposal was supported by the top BLAST hits for the amino acid sequence, as well as the COG, the Pfam, and the PDB. The SignalP and Phobius disagreed on the location of the gene. As such, the proposed annotation is an acl-CoA dehydrogenase. This gene's most likely function is breaking down fatty acids, or metabolizing them.

*Kytococcus sedentarius*08450:

The initial proposed product of these genes by GENI-ACT was a branch-chain amino acid transport protein (AzlD). This was supported by both the Blast and Pfam search. Ksed_08450 is located inside and outside of the membrane because it is a transmembrane. There are three transmembrane helices. It exports amino acids and is known to be involved in resistance to the growth-inhibiting amino acid 4-azaleucine (Belitsky et al., 1997).

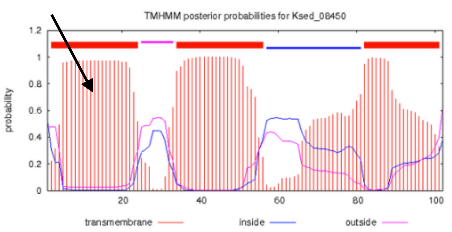

Figure II – The TMHMM results for *Ksed_08450.* The red shows the portions of the protein embedded in the membrane; the black arrow indicates a transmembrane domain. Three transmembrane domains are shown in this graph. The blue and pink lines indicate the portions of the protein that loop inside (blue) and outside (pink) of the membrane.

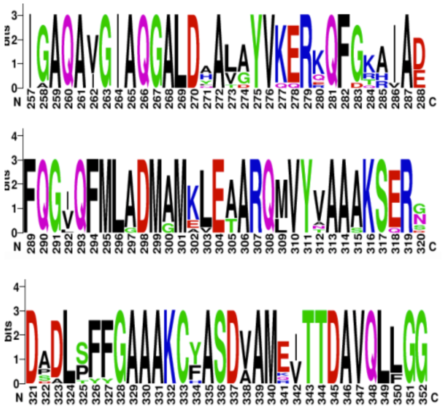

Figure III - A portion of the WebLogo results for *Ksed_08440.* The letters stand for specific amino acids. The colors represent the properties of the amino acids (e.g., polar amino acids are green, acidic amino acids are red). The taller the letter is, and the more room the letter takes up, the more frequently the amino acid is found in that location. This portion of the WebLogo has many highly conserved regions of amino acids.

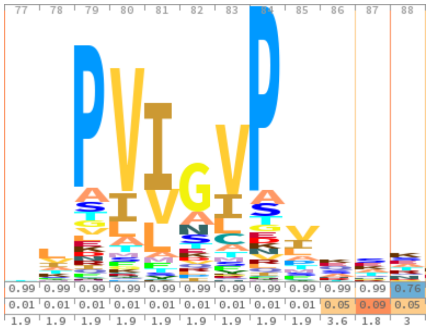

Figure IV – A portion of the HMM logo for *Ksed_08420.* The letters represent existing genes or protein sequences in that are related and can be grouped into families The height and width of the letters is determined by the probability and accuracy of the results. Letters are drawn in different colors, and depends on the amino acid it corresponds to and represents structural or functional similarity.

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 the genes appear to be correctly annotated by the computer database. This conclusion is based on the search results from BLAST, CDD, T-Coffee, and WebLogo for sequence-based similarity data, TIGRfam and Pfam for structure-based similarity data, TMHMM, SignalP and Phobius for cellular localization data.

Gene Locus	Geni-Act Product	Proposed Annotation
Ksed_08420	AIR carboxylase	AIR carboxylase
Ksed_08440	acl-CoA dehydrogenase	acl-CoA dehydrogenase
Ksed_08450	Branched-chain amino acid transport (AzlD)	Branched-chain amino acid transport (AzlD)

References

Belitsky BR, Gustafsson MC, Sonenshein AL, Von Wachenfeldt C; , J Bacteriol 1997;179:5448-5457.: An lrp-like gene of *Bacillus subtilis* involved in branched-chain amino acid transport

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

Acknowledgments

Supported by NSF ITEST Strategies Award Number 1311902