

Annotation of the *Kytococcus sedentarius* Genome from Locus Tags Ksed_09680 to Ksed_09720

Devyn Kelly, Emily Leone, Gavin Luckey, Matthew Pencille 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 four genes from the microorganism *Kytococcus sedentarius* (Ksed 09680 – Ksed_09720) 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 any of the four genes in the group and as such, the genes appear to be correctly annotated by database.

Introduction

Kytococcus sedentarius is a strictly aerobic, non-motile, non-encapsulated, and non-endospore forming gram positive coccoid bacterium, found predominantly in tetrad formation. This organism is classified as a chemoheterotroph, as it requires methionine and several other amino acids for growth. Originally isolated from a microscope slide submerged in sea water in 1944, *Kytococcus sedentarius* grows well in sodium chloride at concentrations less than 10% (w/v).

According to Sims et al. (2009), *Kytococcus sedentarius* is a microorganism of interest for several reasons. This bacterium is a natural source of the oligoketide antibiotics monensin A and monensin B (Sims et al., 2009). *Kytococcus sedentarius* has been implemented as the etiological agent 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 *Micrococccineae*, which has yet to have been thoroughly studied utilizing bioinformatics (Sims et al., 2009).

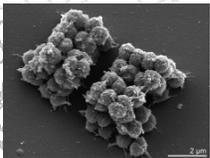


Figure 1 – Image of *Kytococcus sedentarius*.

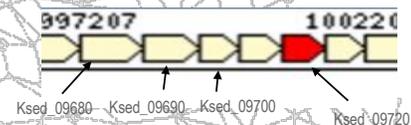


Figure 2 – The locus tags and relative position of the genes under investigation.

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_09680:

The initial proposed product of this gene by GENI-ACT was histidinol-phosphate aminotransferase. This gene product proposal was supported by the top BLAST hits for the amino acid sequence, the presence of well-curated protein functional domains within the amino acid sequence, along with the cellular localization data presented by COG. Histidinol-phosphate aminotransferase is an enzyme that catalyzes a chemical reaction between L-histidinol phosphate and 2-oxoglutarate, while also transferring nitrogenous groups.

Ksed_09690:

The initial proposed product of this gene by GENI-ACT was imidazoleglycerol-phosphate dehydratase/ histidinol-phosphate phosphatase. This gene product proposal was supported by the structure based evidence presented by TIGRFAM, Pfam, and PDB as well as the top BLAST hits for the amino acid sequence. Histidinol phosphatase is an enzyme that catalyzes the chemical reaction, and participates in histidine metabolism. Additionally, in *E. coli*, the enzyme encoded by the gene *hisB* is a fused imidazoleglycerol-phosphate dehydratase and histidinol-phosphate phosphatase.

Ksed_09700:

The initial proposed product of this gene by GENI-ACT was an imidazole glycerol phosphate synthase, glutamine amidotransferase subunit. This gene product proposal was supported by the top BLAST hits for the amino acid sequence along with the cellular localization data presented by COG. Imidazole glycerol phosphate synthase is a key metabolic enzyme that catalyzes the closer of the imidazole ring within histidine biosynthesis.

Ksed_09720:

The initial proposed-product of these genes by GENI-ACT was an Imidazole glycerol phosphate synthase subunit. 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, the cellular location of the amino acid sequences, and the enzymatic function of the amino acid sequences. Imidazole glycerol phosphate synthase is responsible for linking amino acids and biosynthesis.

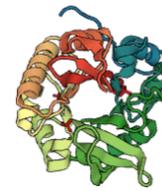


Figure 3 – Imidazole glycerol phosphate synthase protein

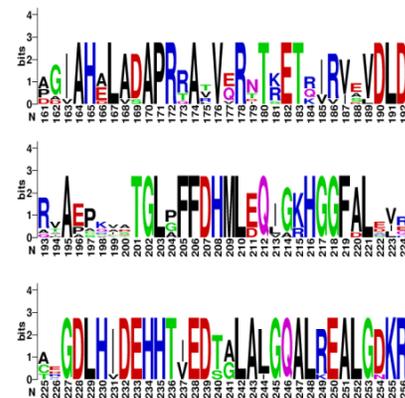


Figure 4 – A portion of the WebLogo for *Ksed_09690* showing a high degree of amino acid conservation found throughout the sequence. Locations showing a single large letter indicate a high degree of conservation, while locations showing multiple smaller letters indicate greater variability. Additionally, the color of the amino acid symbol indicates the properties of each amino acid, for example, polar amino acids are shown as green, basic are blue, acidic are red, and hydrophobic amino acids are shown as black.

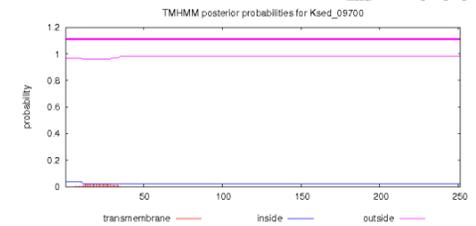


Figure 5 – This graph shows the TMHMM results for *Ksed_09700*. Based on this TMHMM graph, it can be concluded that *Ksed_09700* has no transmembrane helices.

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 Product	Proposed Annotation
09680	Histidinol-phosphate aminotransferase	Histidinol-phosphate aminotransferase
09690	Histidinol-phosphate phosphatase	Histidinol-phosphate phosphatase
09700	Imidazole glycerol phosphate synthase	Imidazole glycerol phosphate synthase
09720	Imidazole glycerol phosphate synthase subunit hisF	Imidazole glycerol phosphate synthase subunit hisF

References

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

Acknowledgments

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

www.buffalo.edu