

# Annotation of the *Kytococcus sedentarius* Genome from Locus Tags

## Ksed\_12650 to Ksed\_12670

Michala Alessandra, Gurlin Kaur, Alyssa Schoenfeldt, and Dennis Bauer

Amherst Central High School and the Western New York Genetics in Research Partnership



University  
at Buffalo

### Abstract

A group of consecutive 3 genes from the microorganism *Kytococcus sedentarius* (Ksed\_12650 – Ksed\_12670) were annotated using the collaborative genome annotation website GENI-ACT. The Genbank proposed gene productname 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 by in their 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).

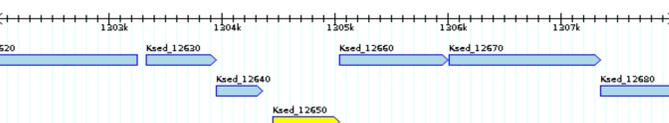


Figure 1. The locus tags and relative position of the genes under investigation in this research

### Methods

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

**Gene Ksed\_12650** is a gene with 591 nucleotide base pairs. Ksed\_12650 is closely related to *Alkaliphilus metalliredigens* QYMF. It codes for the enzyme uracil phosphoribosyltransferase as found in BLAST. This enzyme helps catalyze the reaction needed to create uridine monophosphate (UMP). UMP helps form nucleotides in RNA, specifically pyrimidines. It is found in the cytoplasm as discovered in TMHMM and confirmed in PSORT-B. It is unique to its gene as it has no identical sequences found within the same gene.

**Gene Ksed\_12660**, aspartate carbamoyltransferase (ATcase) is made up of 960 nucleotides and found in *Kytococcus sedentarius* Gene 12660 catalyzes the first step in the pyrimidine biosynthetic pathway. It is likely that gene 12660 is cytoplasmic and since there is no signal peptide the work is done mostly in the cytoplasmic. This supports the idea that this gene is involved in the process of replicating information through cells.

**Gene Ksed\_12670**, Dihydroorotase is a zinc metalloenzyme with a nucleotide length of 1341 and functions throughout the pathway of pyrimidine nucleotide. It catalyzes a reversible interconversion of carbamoyl aspartate and dihydroorotate. The TMHMM results proved that the gene had no transmembrane helices, leading me to believe it could be cytoplasmic. Analysis of the effects of pH, mutant proteins, and alternative substrates on the enzyme-catalyzed reaction, furthers the idea that the gene is located inside the cell and interacts with Ksed\_12650 and Ksed\_12660.

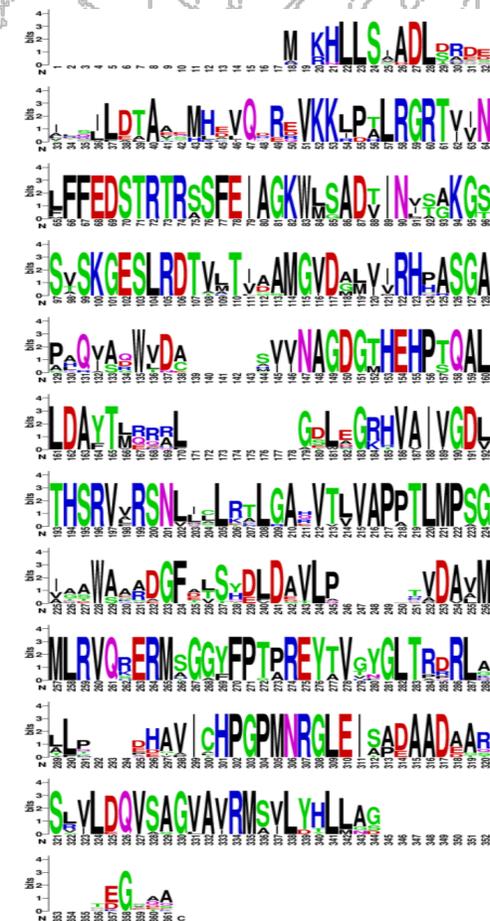


Figure 2 – Ksed\_12660's Weblogo



Figure 3 – Ksed\_12670's Crystal Structure

Modules of the GENI-ACT (<http://www.geni-act.org/>) were used to complete *Kytococcus sedentarius* genome annotation. The modules are described to the left.

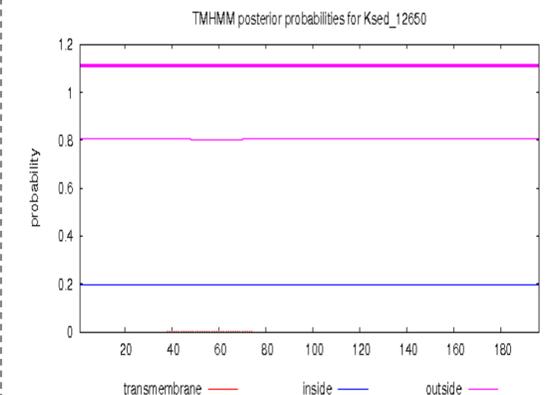


Figure 4 – TMHMM results for Ksed\_12650 shows that the protein lacks any transmembrane helices, suggesting the protein is found in the cytoplasm.

### Conclusion

Pyrimidines ( T and C) pair up with Purines ( A and G) allowing the connection of nucleotides. Nucleotides are the building blocks of DNA and RNA, which gives them their specific function throughout an organism. Gene Ksed\_12670, is a zinc metalloenzyme which functions throughout the pathway of pyrimidine nucleotide. It could catalyze a reversible interconversion of carbamoyl aspartate and dihydroorotate. Gene Ksed\_12660, aspartate carbamoyltransferase (ATcase) catalyzes the first step in the pyrimidine biosynthetic pathway. Gene Ksed\_12650 likely uses the functions of the previous genes to assist in the catalysis that forms RNA. All three of the genes are located in the cytoplasm. Due to the suggested data that these genes are all part of the replication process of a cell it would only be beneficial for all the work to be done within the cell (in the cytoplasm).

### References

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

### Acknowledgments

Special thanks to Dr. Brad Goodner, Dr. Stephen Koury, and Dr. Rama Dey-Rao for their assistance with this project.

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

[www.buffalo.edu](http://www.buffalo.edu)