

Annotation of the *Kytococcus sedentarius* Genome from DNA Coordinates 688202 to 691727

Limariz Lopez, Shea-naya Prim, Zaria Prim, Nicholas Altemoos and Gina O’Kussick

Bennett High School and The Western New York Genetics in Research Partnership

Abstract

A group of consecutive 3 genes from the microorganism *Kytococcus sedentarius* (locus tags Ksed_06940 – Ksed_06960) 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 and potential alternative open reading frames. 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 the computer database. The three genes shared a functionality of phosphate transport as determined by their top BLAST hit names of Phosphate-binding protein PstS 2, Phosphate transport system permease protein PstC 2 and Phosphate ABC transporter, Permease protein PstA, respectively.

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 playground slide submerged in sea water in 1944, *Kytococcus sedentarius* grows well in sodium chloride at concentrations less than 10% (w/v)(Sims et al., 2009).

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 I. Gene neighborhood from IMG-edu. The genes studied are shown in the red box.

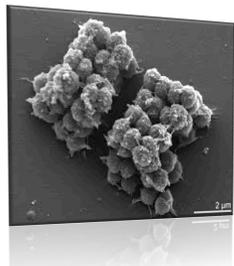


Figure II. Scanning electron Micrograph of *Kytococcus sedentarius* (Sims et al., 2009).

Methods and Materials

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

Kytococcus sedentarius 06950:

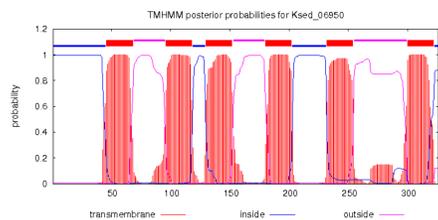


Figure III. TMHMM output predicting the presence of 6 transmembrane helices in Ksed_06950

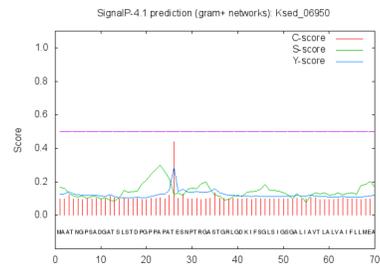
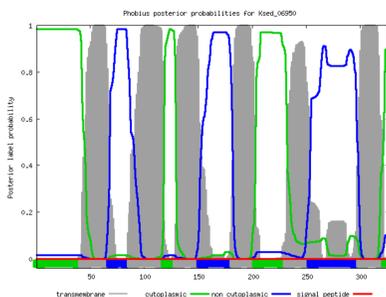


Figure IV. SignalP 4.1 results indicating that Ksed_06950 lacks signal peptide

Measure	Position	Value	Cutoff	signal peptide?
max. C	26	0.438		
max. Y	26	0.288		
max. S	23	0.300		
mean S	1-25	0.155		
D	1-25	0.234	0.450 30	

Figure V. Phobius tool output confirming the lack of a signal peptide and the presence of 6 transmembrane helices in Ksed_06950



Kytococcus sedentarius 06950

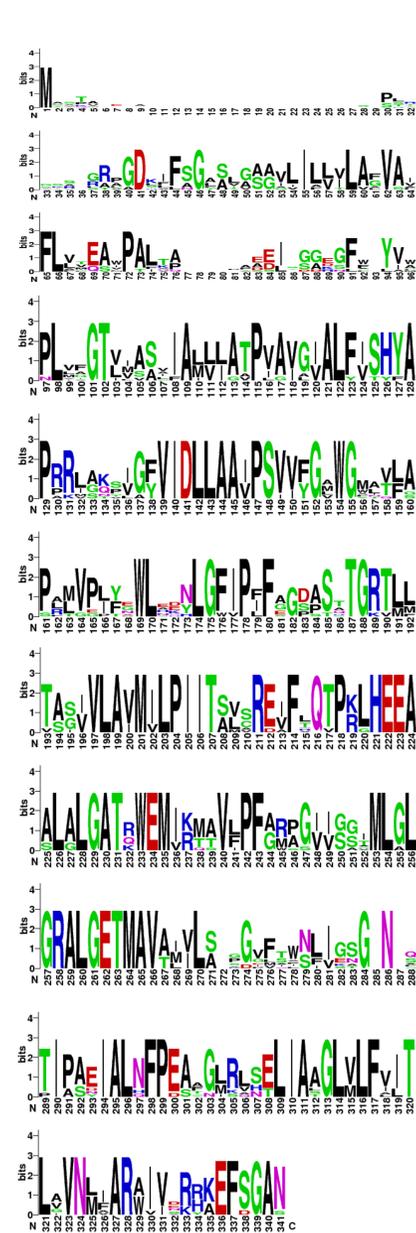


Figure VI – WebLogo of T-Coffee Multiple Sequence Alignment. Repeating domains of hydrophobic amino acids (black letters) can be seen, which correspond to the transmembrane helical domains.

Ksed_06950 Summary

Ksed_06950 appears to have several well conserved amino acid domains, as shown in Figure VI. A Phobius and TMHMM output shows the presence of 6 transmembrane helices. Psortb predicts a cytoplasmic membrane score of 10.00. All evidence suggest this is a cytoplasmic membrane protein.

Ksed_06960 Summary

Ksed_06960 also shows the presence of 8 transmembrane helices by TMHMM and Phobius (Figures VII and VII) and has a cytoplasmic membrane score of 10.00 in Psortb. It is also therefore likely to be a cytoplasmic membrane protein

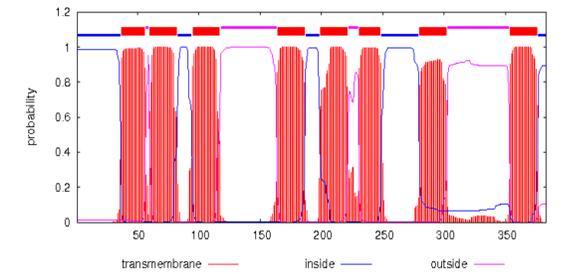


Figure VII - TMHMM tool output for Ksed_06960 showing the presence of 8 transmembrane helices.

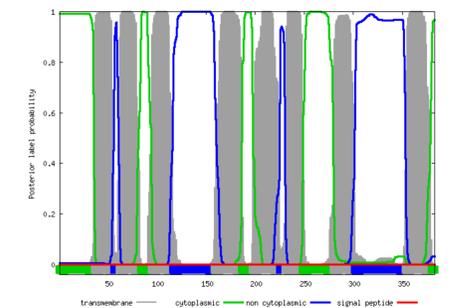


Figure VIII – Phobius tool output confirming the presence of 8 transmembrane helices in Ksed_06960. No signal peptide is indicated.

Conclusion

Gene Locus Tag	Top Blast Species/Gene Product Name
06940	Phosphate-binding protein PstS 2 <i>Mycobacterium tuberculosis</i> CDC1551
06950	Phosphate transport system permease protein PstC 2 <i>Mycobacterium bovis</i> AF2122/97
06960	Phosphate ABC transporter, Permease protein PstA <i>Mycobacterium leprae</i> TN

References

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 Dr. Stephen Koury Research Associate Professor, University at Buffalo Dr. Rama Dey-Rao Clinical Assistant Professor, University at Buffalo