

Annotation of the *Kytococcus sedentarius* Genome from DNA Coordinates 679771 to 682339

Lily Freeman-Striegel, Amandip Deo, Cole Gleichenhaus and Dennis Bauer

Department of Biological Studies @ Amherst Central High School and the Western New York Genetics in Research Partnership

Abstract

A group of 3 consecutive genes from the microorganism *Kytococcus sedentarius* (Ksed_06870 – Ksed_06890) 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, 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 the r database.

Introduction

Kytococcus sedentarius is a gram positive coccoid bacterium, found predominantly in tetrads, irregular clusters and cubicle packets of 8. Optimal growth temperature is from 25°C to 37°C. The genome has been sequenced and contains 2,785,024 base pairs, 2,639 protein coding and 64 RNA genes. This organism is classified as a chemoheterotroph, meaning it requires methionine and several other amino acids for growth. Originally isolated from a microscope slide submerged in seawater in 1944, *Kytococcus sedentarius* grows well in sodium chloride.

This gene was studied in order to determine the full genome of itself and the possibility of further exploring other genomes. Throughout the weeks that we studied this gene, we've expanded our knowledge on genomics and learned how to use many important and interesting tools and databases. By us further researching this, we've come closer to preventing mild cases of pneumonia and extreme cases of pitted keratolysis. It is essential to annotate genomes to further our understanding of biology.

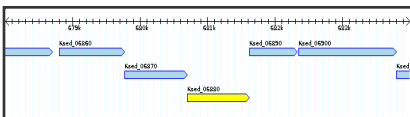


Figure 1. Gene neighborhood containing the locus tags analyzed in this study: Ksed_06870, Ksed_06880, Ksed_06890

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?
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

Kytococcus Sedentarius 06870
 We determined that this gene most likely codes for a membrane protein. We decided this because of the TMHMM (Transmembrane Helices; Hidden Markov Mode) results showing evidence of multiple helices, or folds in the protein that cross the cellular membrane. This concludes that ksed 06870 has the highest probability of being a membrane protein.
Kytococcus Sedentarius 06880
 Through extensive research, we found out that this portion of the genome binds adenosine triphosphate. This gene is known as an ABC transporter. It does not have any transmembrane helices, and it is a cytoplasmic protein. We found this out by doing the TMHMM test, which showed results that were very well conserved.
Kytococcus Sedentarius 06890
 This gene was very interesting in that it showed similarities to other helix-displaying genes, but did not show any evidence of having helices in the cell membrane. Though this gene has no helices in the cellular membrane, we predicted that this gene is probably an ABC transporter; Helices may be present in the mitochondrial membrane which would support the fact that this gene is very similar to other helix-displaying membrane proteins.

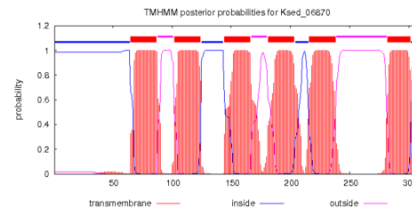
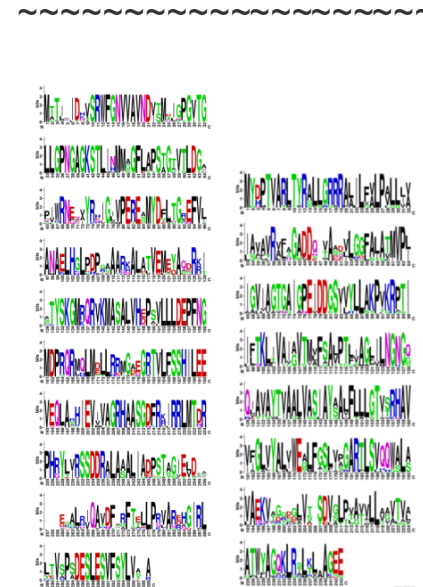


Figure 2. TMHMM results for Ksed_06870 showing the presence of 6 transmembrane helices.



Eye-appealing WebLogo biological illustrations showing significant conservation of sequences.



Foot affected by pitted keratolysis caused by *Kytococcus sedentarius*

Conclusion

This chart shows our overall conclusions about our specific genes in *Kytococcus Sedentarius*

Gene Location	Hypothesized Function
>kсед_06870	Transporter protein on cellular membrane
>kсед_06880	ABC Transporter
>kсед_06890	ABC Transporter

References

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

Acknowledgments

Special thanks to Dr. Stephen Koury and Dr. Rama Dey-Rao for their assistance with this project. Supported by NSF ITTEST Strategies Award Number 1311902