

Annotation of the *Kytococcus sedentarius* Genome from Locus Tags

Ksed_12560 to Ksed_12580

Owen Boccabella, Madeline Buisch, Ian Leising and Dr. Paul Weber

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



University at Buffalo

Abstract

A group of 3 consecutive genes from the microorganism *Kytococcus sedentarius* (Ksed_12560 – Ksed_12580) 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 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 seawater 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).

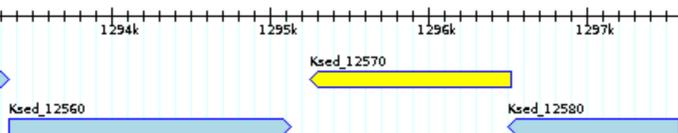


Figure 1. The locus tags and relative position of the genes under investigation in this research

Methods

Modules 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

Ksed_12560:

The initial proposed product of this gene by GENI-ACT was a shikimate kinase/3-dehydroquinate synthase. 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, the transmembrane topography of the amino acid sequence, and the cellular location of the amino acid sequence. As such, the proposed annotation is a sodium/proline symporter.

Ksed_12570:

The initial proposed product of this gene by GENI-ACT was an oligopeptide transport ATP-binding protein. This gene product proposal was supported by the top BLAST hits for the amino acid sequence, also the CDD top hits supported the proposed protein as there was a result of dipeptide transporter ATP-binding subunit from the amino acid sequence. Ksed_12570 was thought to be cytoplasmic by Figure 4. A later test (Figure 4) showed that did not have transmembrane helixes, suggesting Ksed_12570 is a cytoplasmic protein or that is transported to the outside of the cell.

Ksed_12580:

The initial proposed product of this gene by GENI-ACT was an oligopeptide/dipeptide ABC transporter, ATP-binding protein. This gene product proposal was supported by the top BLAST hits for the amino acid sequences, the sequence alignment with proteins of the suggested protein family, protein structure, the top BLAST hits for paralogs, and the function of related proteins. This gene product proposal is also supported by Figure 5 Due to this, the proposed annotation is an ABC transporter ATP-binding protein.

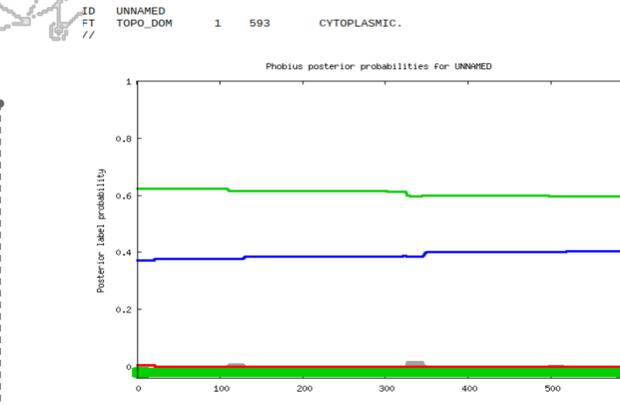


Figure 2 – Phobius results for Ksed_12560, showing that it lacks both a signal peptide and transmembrane helixes, suggesting it is located in the cytoplasm.

Tree Rendering results

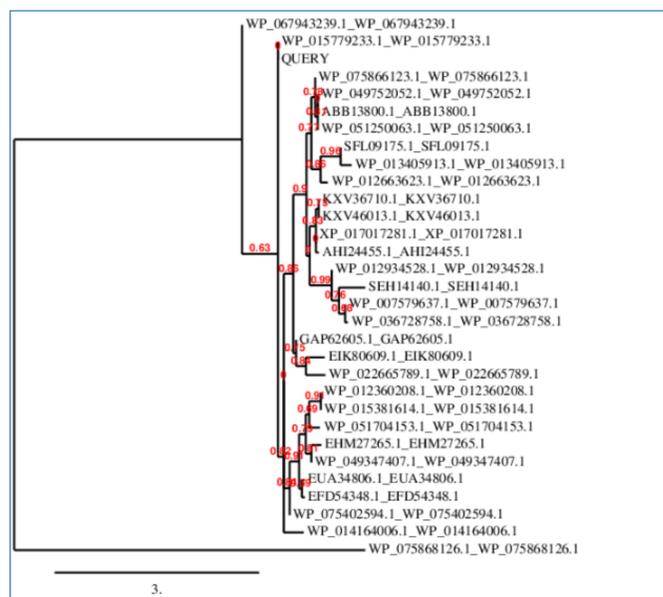


Figure 3-Ksed_12560 phylogenetic tree displays related proteins with a similar function to Ksed_12560.

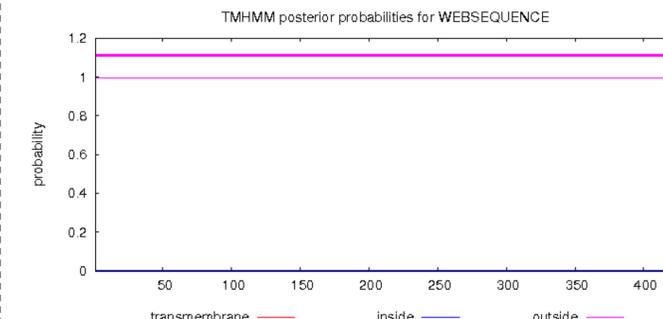


Figure 4 – Ksed_12570 is located outside of the cell or in the cytoplasm.

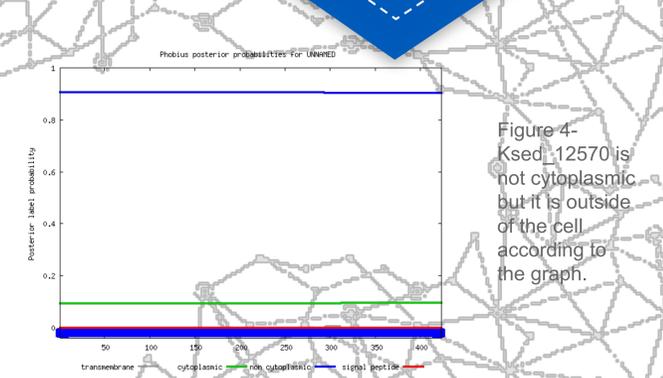


Figure 4- Ksed_12570 is not cytoplasmic but it is outside of the cell according to the graph.

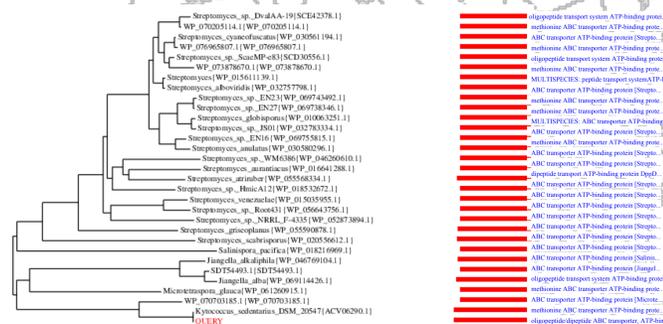


Figure 5 – Ksed_12580 phylogenetic tree displaying proteins from closely related organisms. This suggests the possible function of the protein is an ABC transporter ATP binding protein.

Conclusion

Ksed_12560: Originally, Ksed-12560 was proposed to be shikimate kinase/3-dehydroquinate synthase and I agree with this proposition. This strand of ksed could possibly eat flesh, due to the fact that it is located in the cytoplasm of the cell, as seen in Figure 2.

Ksed_12570: Theoretically Ksed_12570 could take part in flesh eating since it seemed to be outside the cell since it did not test positive for cytoplasmic. A possible function of Ksed_12570: is that it might take part in the eating of flesh since it could be outside of the cell according to figure 4.

Ksed_12580: This gene was proposed to be an oligopeptide transport ATP-binding protein, but the proposed annotation is an ABC transporter ATP-binding protein

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