Annotation of the *Kytococcus sedentarius* Genome from Locus Tags Ksed_12590 to Ksed_12610)

Serena Bangia, Sophia Lawrence, Sophie Smith and Dr. Paul Weber
Amherst Central High School, Buffalo, 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_12590 – Ksed_12610) 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 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 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 *Micrococcineae*, which has yet to have been thoroughly studied utilizing bioinformatics (Sims et al., 2009).

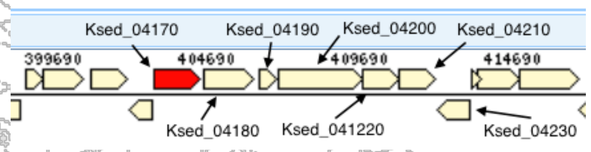

Figure 1. The locus tags and relative position of the genes under investigation in this research

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_12590

It seems there may be 6 transmembrane helixes in this protein as shown by the TMHMM results in Figure 2. Since the top blast hit for this protein seems to be the Oligopeptide transport system permease protein it would make sense since this is a permease that the protein is associated with the membrane. Figure 3 shows Phobius results confirming the presence of 6 transmembrane helixes and the lack of a signal peptide.

Ksed_12600

The Phobius results shown if Figure 4 suggest this protein passes through the membrane six times, strongly suggesting it is a transmembrane protein. Passing through the membrane fewer times indicates it may not transport anything through the membrane. Since it passes through multiple times, it can deliver things to and from the cytoplasm. This finding is in agreement with the top BLAST hit name of ABC transporter permease, which would be expected reside within the cell membrane.

Ksed_12610

Phobius results (Figure 5) indicate the presence of a signal peptide in Ksed_12610. The graph has a significant spike in the amino acid region around 20. This spike conveys there is a signal peptide which spikes from 0-1.0 posterior label probability. As the signal peptide decreases from 1.0-0 the non cytoplasmic appears to increase around amino acid region 20 as well from 0-1.0 posterior label probability. The presence of a signal peptide is in agreement with the top BLAST hit name of ABC transporter substrate-binding protein, which would be expected to be secreted to the outside of the cell membrane.

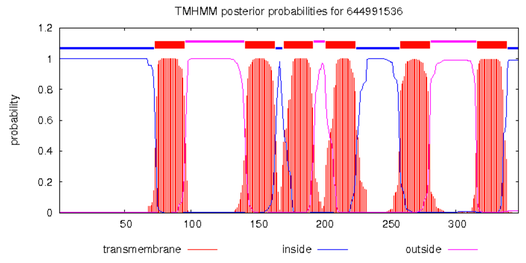

Figure 2. TMHMM results for Ksed_12590 in which 6 transmembrane helixes are predicted

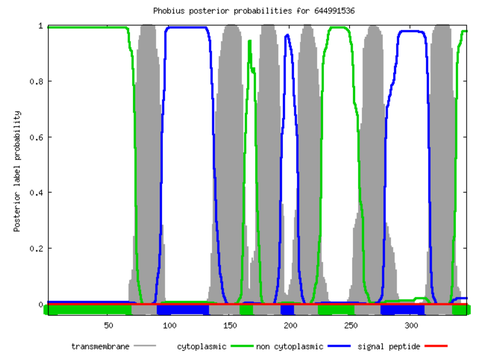

Figure 3. Phobius results for Ksed_12590 confirming that Ksed_12590 has 6 transmembrane helixes and that if further lacks a signal peptide.

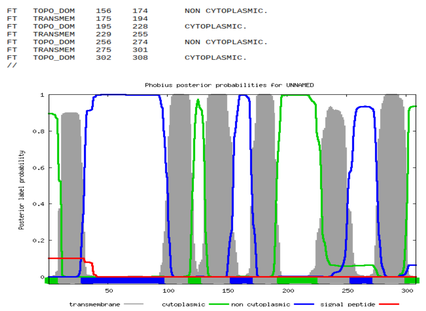

Figure 4 – Ksed_12600 graph indicates that Ksed_12600 is transmembrane protein , travelling in and out of the cytoplasm.

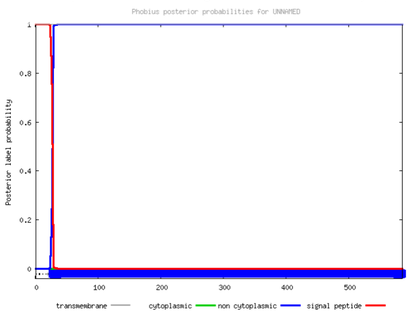

Figure 5 – Phobius results for Ksed_12610 showing the presence of a signal peptide (red line)

Ksed_12590
The Phobius and TMHMM results clearly suggest that this could be a transmembrane receptor because it has three areas inside the membrane, shown in Figure 2. It seems it has enough transmembrane domains to be a transport protein like 12600.

Ksed_12600
The Phobius results indicate that this could be a transport protein because it enters and exits the membrane multiple times. It would seem to function in concert with Ksed_12590.

Ksed_12610
The Phobius results suggests that there is a signal peptide occurring around amino acid region 20, and that there are no transmembrane helixes. This means means Ksed_12610 is translocated across the membrane to the exterior of the cell rather than being embedded in the membrane. As the top BLAST hit is a ABC transporter substrate binding protein. It could help other proteins transfer across the membrane by way of the permeases encoded by Ksed_12590 and Ksed_12600.

Conclusions

Ksed_12590- Oligopeptide transport system permease protein AppC- There is no evidence currently in this research project to suggest that this is definitely an oligopeptide system, however each individual protein could oligomerize to form a circle and each go through the membrane twice.

Ksed_12600- The suggested protein was the Dipeptide transport system permease protein. Evidence supported this suggestion, because of the amount of transmembrane areas shows that it could be a transport system.

Ksed_12610- The data suggests the protein is an oligopeptide binding protein. The protein involves a signal peptide which interacts with the membrane.

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