

Annotation of the *Kytococcus sedentarius* Genome from DNA Coordinates 08030 to 08060

Xiaoyu Li, Jessica Meng, Gouthamrajan Nadarajan, Adityanarayanan Nair, Kathleen Kramer

Williamsville East High School and the Western New York Genetics in Research Partnership

Abstract

4 genes from the microorganism *Kytococcus sedentarius* (Ksed_08030 – Ksed_08060) were annotated using the computer program Genomics Education National Initiative Annotation Collaboration Toolkit (GENI-ACT). The proposed gene product for each gene was manually annotated with GENI-ACT test modules including Cellular Localized Data, Structure-Based Evidence, and Sequence-Based Similarity Data. For all 4 genes, the Genbank proposed gene product did not differ from the proposed annotation. Genes 08030 and 08060 were found to code for transmembrane transport proteins; gene 08040 codes for a transcription regulator; gene 08050 codes for an enzyme that is involved in DNA cleavage and cell division.

Introduction

Kytococcus sedentarius is a skin bacterium that is usually harmless but can cause serious infections such as valve endocarditis, hemorrhagic pneumonia, and pitted keratolysis (Sims et al., 2009). However, it is also a producer of the oligopeptide antibiotic monensin A, which is widely used in the poultry and cattle industries (Lowicki et al., 2013). This organism must be researched to further clarify its virulence and discover what benefits it may offer to humans.

Kytococcus sedentarius is a nonmotile, Gram-positive bacterium, originally found in marine environments. Gram positive bacterium. It is strictly aerobic, and can only grow when amino acids are provided on the medium. It is a producer of the oligopeptide antibiotics monensin A and B, but DNA analysis revealed no close relationship to *Streptomyces cinnamonensis*, a typical monensin producer (Pospisil et al., 1998). It is a part of the family Dermacoccaceae, which is not very well studied (Sims et al., 2009).

The current study of *Kytococcus sedentarius* focuses on using the computer annotating program, Genomics Education National Initiative Annotation Collaboration Toolkit (GENI-ACT), to manually annotate the genes of the bacteria. Through this, the gene products can be confirmed or be marked for further investigation based on data obtained from the GENI-ACT tests. Some of the tests performed include TMHMM, a program that predicts the probability of transmembrane helices; Phobius, which predicts the probability of transmembrane helices, signal peptides, cytoplasmic and non-cytoplasmic proteins; and MetaCyc, which shows maps of enzymatic pathways in which the gene product may participate.

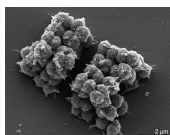
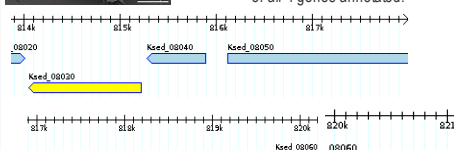


Figure 1. Scanning electron micrograph of *K. sedentarius* (Sims, et al., 2009)

Figure 2. Gene neighborhoods of all 4 genes annotated.



Methods and Materials

Modules of the GENI-ACT (<http://www.geni-act.org/>) were used to complete the *K. sedentarius* genome annotation.

The modules are described below:

Modules	Activities	Questions Investigated
Basic Information	DNA Coordinates and Sequence, Protein Sequence	What is the sequence of gene and protein? Where is it located in the genome?
Sequence-Based Similarity	Blast, CDD, Tcoffee, WebLogo	Is my sequence similar to other sequences in Genbank?
Cellular Localization Data	Gram Stain, TMHMM, SignalP, PSORT-b, Phobius	Is my protein in the cytoplasm, secreted or embedded in the membrane?
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?
Structure-based Evidence	TIGRFam, Pfam, PDB	Are there functional domains in my protein?
Enzymatic Function	KEGG, MetaCyc, E.C. Number	In what process does my protein take part?
Duplication and Degradation	Paralog, Pseudogene	Are there other forms of my gene in the bacterium? Is my gene functional?
Evidence for Horizontal Gene Transfer	Phylogenetic tree	Has my gene co-evolved with other genes in the genome?

Results

08030: The gene product for Ksed_08030 is an arabinose efflux permease that functions in cellular transport of the monosaccharide arabinose. This product is supported by investigations in the Cellular Localized Data module of GENI-ACT. TMHMM and Phobius both predicted 12 transmembrane helices, and Phobius predicted a very low chance of the presence of a signal peptide, so the protein is highly unlikely to be secreted from the cell. This leads to the conclusion that the gene product is an integral transmembrane transport protein.

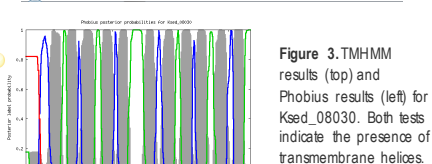
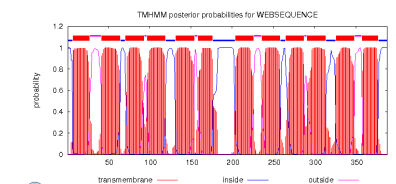


Figure 3. TMHMM results (top) and Phobius results (left) for Ksed_08030. Both tests indicate the presence of transmembrane helices.

08040: The gene product of Ksed_08040 is a transcriptional regulator, meaning it is involved in the conversion from DNA to RNA. According to Pfam, it is part of the TetR family, and its Pfam domain name is Helix Turn Helix which is characterized by a structure that is capable of binding to DNA. In Module 2 Sequence-based Similarity Data, the top hit on BLAST was NR transcriptional regulator [*Gordonia otitidis*] with an e-value of 9e-95, and the second hit was NR transcriptional regulator [*Langella gansuensis*] with an e-value of 5e-86. In addition, according to the investigations in Module 3, Cellular Localization Data, on Phobius, PSORT-B, and LipoP, it is most likely a cytoplasmic protein, and according to TMHMM, there are 0 likely transmembrane regions; all of these investigations further support that it is a transcriptional regulator that remains in the cytoplasm of the cell. The WebLogo of the gene indicates that residues 1-26 are mostly well-conserved non-polar amino acids, 27-62 is a mostly well-conserved polar region, and 65-200 is mostly moderately or poorly conserved.

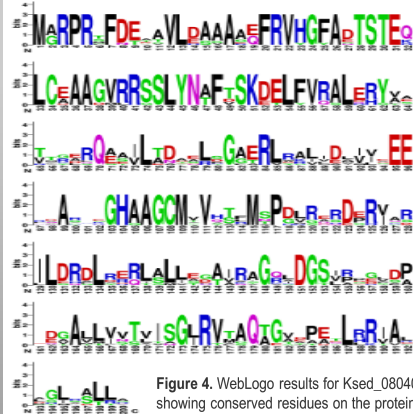
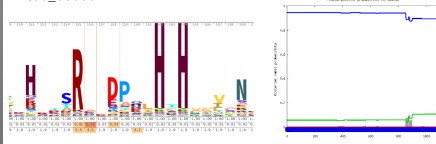


Figure 4. WebLogo results for Ksed_08040, showing conserved residues on the protein.

08050: The gene product for Ksed_08050 is TrwC relaxase, which is an enzyme that has DNA cleavage and strand transfer activities during cell division. This evidence is supported through the results of Tigrfam, PSORT B, and the HMM logo. According to Tigrfam, we found out that this enzyme acts on plasmid R388 conjugation and cleavage. PSORT B results confirmed that this gene product is located in the cytoplasm which further supports the conclusion because a bacterium doesn't have a nucleus. The HMM logo showed us that the gene is well conserved and thus strengthens our prediction.

Figure 5. The HMM logo (left) and Phobius results (right) for Ksed_08050.



08060: The product of Gene sequence Ksed_08060 is an arabinose efflux permease family protein, which functions in the transport of arabinose, a sugar of the pentose class. This is proven by the PSORT-B test because only the cytoplasmic membrane score was above zero. The TMHMM test also predicted this as having twelve helices. The signal peptide probability was listed as not likely as well and that indicates that it is a transmembrane protein and not likely to be secreted by the cell. The PFAM investigation also assists in proving the function of the gene. Its PFAM name is listed as Major Facilitator Superfamily. The major facilitator superfamily (MFS) is a class of membrane transport proteins that facilitate movement of small solutes across cell membranes. This indicates that the following gene product is a transporter protein reinforcing the fact that it transports arabinose across the membrane.

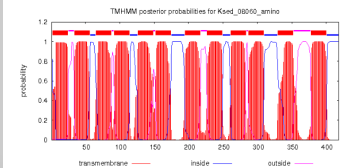


Figure 6. The TMHMM results for Ksed_08060, indicating that the gene product is a transmembrane protein.

Conclusion

The Genbank proposed gene products were all confirmed:

Gene Locus	Genbank Product	Proposed Annotation
08030	Arabinose efflux permease	Arabinose efflux permease
08040	Transcriptional regulator	Transcriptional regulator
08050	TrwC relaxase	TrwC relaxase
08060	Arabinose efflux permease	Arabinose efflux permease

References

Lowicki et al. (2013). Structure and Antimicrobial Properties of Monensin A and Its Derivatives: Summary of the Achievements. *BioMed Research International* Volume 2013 (2013), Article ID 742149

Pospisil, et al. (1998). *Kytococcus sedentarius* (formerly *Micrococcus sedentarius*) and *Demococcus nishinomyensis* (formerly *Micrococcus nishinomyensis*) produce monensins, typical *Streptomyces cinnamonensis* metabolites. *Canadian Journal of Microbiology*, 44: 1007-1011.

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

Acknowledgments

Special thanks to Dr. Rama Dey-Rao and Dr. Stephen Koury from University at Buffalo. Supported by NSF ITEST Strategies Award Number 1311902