Annotation of the *Kytococcus sedentarius* Genome from Locus Tags Ksed_08620 to Ksed_08660

Meghan Gallivan*, Katelyn Whitford*, Kristen Wielgasz*, Dariann Russo and Cheryl Smith

Silver Creek High School and the Western New York Genetics in Research Partnership

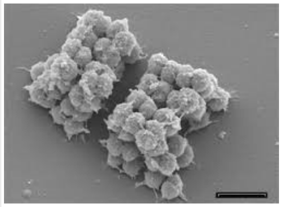
Abstract

The purpose of the Geni-Act bioinformatics research was to prove the computer programs used correctly analyzed genes from *Kytococcus sedentarius*. The four genes researched were locus tags Ksed_08620, Ksed_08630, Ksed_08650, and Ksed_08660. Geni-Act was connected to 9 different modules designed to analyze the genes using the basic amino acid sequences, sequence similarities, structure-based evidence, cellular localization and open reading frame analysis. Results for the gene products of all genes were as expected from the Genbank. There were some results which differed from what was predicted by the databases and computer programs indicating the importance of manual annotation.

Introduction

Kytococcus sedentarius is a gram positive organism that is found mainly on human skin, though it was originally located in a marine environment. It was discovered 1944 in a microscope slide of sea water. It was originally known as *Micrococcus sedentarius*, but was changed after further research. *Kytococcus sedentarius* is an organism of interest because of its being a member of the *Dermacoccaceae* family which has been a poorly studied family. The complete genome of *K. sedentarius* has not been previously manually annotated. The purpose of this study was to manually annotate the genome of the bacterium and to describe the function of the gene products of *K. sedentarius*.

Kytococcus sedentarius can be found in tetrads, irregular clusters, and cubical packets of eight. The bacterium is mainly isolated and grows best at a temperature between 25-37 ºC. *K. sedentarius* contains two enzymes that have been found to degrade keratin, which causes the condition pitted keratolysis in humans. It also plays a role as an opportunistic pathogen causing infections such as hemorrhagic pneumonia and valve endocarditis. Also, the bacterium is known for producing oliogoketide antibiotics, monensin A and monensin B, which are used in the treatment of illness.



Kytococcus sedentarius

Pitted Keratolysis (2016)

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 08620:

The initial proposed result of the gene product from the research done through GENI-ACT was phosphoribosyltransferase. This initial proposal was supported by the top hit in BLAST through the amino acid sequence. TMHMM and PSORT-B support the fact that the enzyme is in the cytoplasm. Based on these results, the gene product of Ksed_08620 was confirmed as amidophosphoribosyltransferase , an enzyme protein in the cytoplasm.

Kytococcus sedentarius 08630:

The expected result was that the gene Ksed_08630 was to be a cytoplasmic protein that did not have any transmembrane helices and was not an enzymatically functioning protein. Evidence of these facts were from TMHMM, KEGG, and PSORT-B. There was also evidence given that Ksed_08630 is a ribosomal protein. However, there was a small amount of evidence shown in the alternative reading frame that suggests that the computer analyzed the Ksed_08630 start codon incorrectly. The computer proposed that start codon to be at the DNA coordinates 877860 however there was no Shine-Delgarno sequence shown before those coordinates. In return, the newly proposed coordinates of the start codon should be 877777.

Kytococcus sedentarius 08650:

The initial proposed result of the gene product from the research done through GENI-ACT is a cytoplasmic acetyltransferase enzyme. This proposal is supported by the top hits from BLAST for the amino acid sequence, as well as the presence of well-curated functional domains with evidence on information found in Pfam and PDB. TMHMM, SignalP, and PSORT-B all showed results the location of the protein would be in the cytoplasm therefore the hypothesis is that the protein is cytoplasmic. Based on this, the hypothesis is that Ksed_08650 codes for a cytoplasmic acetyltransferase. The function of acetyltransferase is to transport acetyl groups in making fatty acids.

Kytococcus sedentarius 08660:

The initial protein product for this gene is a Protein translocase subunit Sec A. The function of this protein is to transport proteins across and through the cell membrane. The protein becomes active after it hydrolyzes a molecule of ATP. This protein is also involved in intracellular trafficking, secretion, and vesicular transport. This is based on evidence found through the program BLAST. The results from the program Weblogo shows that this protein is overall highly conserved across a variety of organisms. The results from TMHMM, SignalP, and PSORTb show that the location of this protein is cytoplasmic. This is logical since Sec translocase proteins exist in the cytoplasmic membrane of bacterial organisms (Natale et al, 2008).

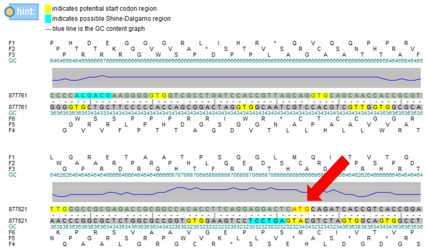


Figure I: *Kytococcus sedentarius* 08630: The original proposed DNA coordinates at 877860. There is no Shine-Delgarno sequence located before the start codon at 877860 (red arrow)

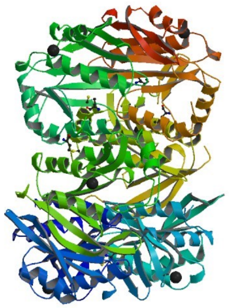


Figure II : *Kytococcus sedentarius*_08650: A crystalline protein structure of acetyltransferase.

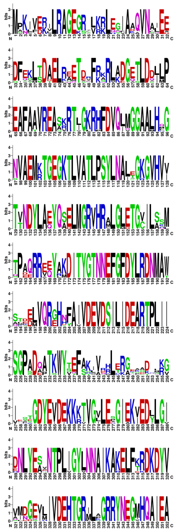


Figure IV.

Weblogo for Ksed_08660 indicates how conserved the protein is. This shows that Protein translocase subunit Sec A is overall highly conserved across a variety of organisms.

Conclusion

The GENI-ACT proposed gene products did not differ significantly from the proposed gene annotation for each of the genes. The genes appear to be correctly annotated by the computer database.

Gene Locus	Proposed Gene Products
Ksed_08620	amidophosphoribosyltransferase
Ksed_08630	ribosomal protein
Ksed_08650	acetyltransferase
Ksed_08660	Protein translocase subunit Sec A

References

"Pitted Keratolysis Pictures." *PittedKeratolysis*. PittedKeratolysis.com, 13 Mar. 2012 Web. 12 May 2016

Chaudhary, D., and S.N. Finkle. "Peritoneal Dialysis International." *Peritoneal Dialysis-Associated Peritonitis Due to Kytococcus Sedentarius*. Multimed Inc, 2016. Web. 11 May 2016

Heinz, D. W., W A. Baase, F.W. Dahlquist, and B.W. Mathews. "1NSL." *RCSB PDB*. National Science Foundation, 1993. Web. 11 May 2016

Hornick, Leila. "Home - Kytococcus Sedentarius DSM 20547." *Home- Kytococcus Sedentarius DSM 20547.* U.S Department of Energy, 21 Apr. 2016. Web 11 May 2016

Sec- and Tat-mediated Protein Secretion across the Bacterial Cytoplasmic Membrane--distinct Translocases and Mechanisms." *Http://www.nlm.nih.gov/*. N.p., n.d. Web.

Torvalds, Linus. "Virtual Unknown(tm) Microbiology Internet Edition 2012" *Virtual Unknown (tm) Microbiology Internet Edition 2012*. Intuitive Systems Inc, 2011. Web 11 May 2016

Acknowledgments

Supported by NSF ITEST Strategies Award Number 1311902