University at Buffalo The State University of New York | REACHING OTHERS

Annotation of the *Kytococcus sedentarius* Genome from DNA Coordinates 403939 to 414240

Alicia Majkut, Mistie Keller, Olivia Slotman, Emily Guest, Dakota Glashauser, Ryan Kostusiak and Jeff Salerno

Lake Shore High School and the Western New York Genetics in Research Partnership

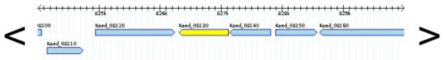
Abstract

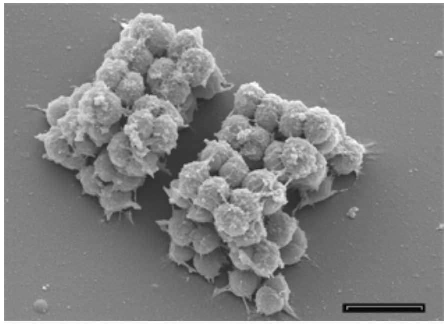
A group of 6 consecutive genes from the microorganism *Kytococcus sedentarius* (Ksed_08120 – Ksed_08180) were annotated using the collaborative genome annotation website GENI-ACT. The Genbank proposed gene product name for each gene was assessed in terms of the general genomic information, amino acid 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 and the possibility of horizontal gene transfer. 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 ther 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 *Micrococcineae*, which has yet to have been thoroughly studied utilizing bioinformatics (Sims et al., 2009).





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?

Results

Kytococcus sedentarius 08120:

The initial proposed product of this gene by GENI-ACT was a Thymidine phosphorylase. 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. The purpose of this gene is to metabolize pyrimidine. Thymidine phosphorylase is overexpressed in various tumors and plays an important role in tumor growth.

Kytococcus sedentarius 08130:

The initial proposed product of this gene by GENI-ACT was a 2-deoxyribose 5-phosphate aldolase. This gene product proposal was supported by the top BLAST hits for the amino acid sequence, the presence of well-curated functional domains within the amino acid sequence, the cellular location of the amino acid sequence, and the enzymatic function of the amino acid sequence. As such, the function of this gene is to catalyze a reversible reaction from acetaldehyde and D-glyceraldehyde 3-phosphate to generate 2-deoxy-D-ribose 5-phosphate. It is also in the pentose phosphate pathway(glycolysis) and ATP production.

Kytococcus sedentarius 08140

The initial proposed product of these genes by GENI-ACT was an ATP-binding protein. This gene product proposal was supported by the top BLAST hits for the amino acid sequences, the presence of well-curated functional domains within the amino acid sequences, the cellular location of the amino acid sequences, and the enzymatic function of the amino acid sequences. The purpose of this gene is to metabolize and to transfer phosphorus-containing groups.

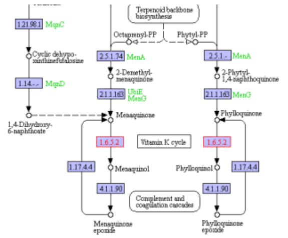


Figure I – *Kytococcus sedentarius* 08180 pathway map from KEGG, which defines enzymatic pathways, subcellular structures, and a number of other modular or sequential functions and processes in terms of maps that visualize the pathway or module and the genes involved in them.

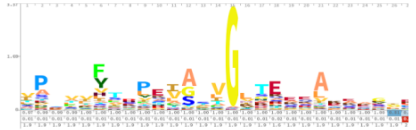


Figure II – *Kytococcus sedentarius* 08180 HMM Logo. This provides the a quick overview of the features of a profile HMM while conserving as much information as possible.

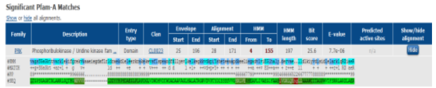


Figure III – *Kytococcus sedentarius* 08140 Pfam. Pfam is a collection of hidden Markov models (HMMs) Pfam can be used to identify protein domains found in a query sequence.

Figure IV - *Kytococcus sedentarius* 08160 Phylogenetic Tree: Rooted phylogenetic trees show relationships with a timeline. Often the length of the lines in a rooted phylogenetic tree is related to time. By contrast, an unrooted phylogenetic tree shows a snapshot of relatedness without a timeline. An unrooted phylogenetic tree clusters similarities together but does not show an ancestral lineage. This tree of the gene 08160 portrays that Kytococcus sedentarius shares this gene with many other organisms of the same order with minor differences but the same function.

Kytococcus sedentarius 08160:

The initial proposed product of this gene by GENI-ACT was a phosphomannomutase. This gene product proposal was supported by nearly all of the separate programs. It was found to be a protein that was always located inside of the cytoplasm and converted α-D-mannose 1-phosphate to D-mannopyranose 6-phosphate. Genes of similar function are found in many other organisms.

Kytococcus sedentarius 08170:

The initial proposed product of this gene by GENI-ACT was a Purine nucleoside phosphorylase. This gene product proposal was supported by the top BLAST hits for the amino acid sequence, the presence of well-curated functional domains within the amino acid sequence, the transmembrane topography of the amino acid sequence, the cellular location of the amino acid sequence, and the enzymatic function of the amino acid sequence. This gene product proposal is involved in purine metabolism. It is an enzyme in humans that is encoded by the NP gene.

Kytococcus sedentarius 08180:

Initially this proposed product was a component of 2-oxoglutarate dehydrogenase complex. Dehydrogenase is an enzyme belonging to the group of oxidoreductases that oxidizes a substrate by a reduction reaction that removes one or more hydrogens from a substrate to an electron acceptor, usually $NAD^+/NADP^+$ or a flavin coenzyme such as FAD or FMN. This gene product proposal was supported by the top BLAST hits for the amino acid sequence. The presence of related enzymes have been observed acting on a sulfur group of donors with NAD+ or NADP+ as acceptor.

Conclusion

The GENI-ACT proposed gene product 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 the computer database.

Gene Locus Tag	Geni-Act Gene Product Name	Cellular Localization
08120	Thymidine phosphorylase	Cytoplasmic
08130	2-deoxyribose 5-phosphate aldolase	Cytoplasmic
08140	ATP-binding protein	Cytoplasmic
08160	Phosphomannomutase	Cytoplasmic
08170	Purine nucleoside phosphorylase	Cytoplasmic
08180	Dehydrogenase	Cytoplasmic

References

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

Acknowledgments

This work was supported by NSF ITEST Strategies Award Number 1311902.

Dr. Stephen Koury Research Associate Professor, University at Buffalo