Annotation of the *Kytococcus sedentarius* Genome from Locus Tags Ksed_09150 to Ksed_09190

Victoria Medley, Duska Samardzija, Alexis Smith, Ari Yarbourough, and Christine Lamont

Buffalo Academy of Visual and Performing Arts and the Western New York Genetics in Research Partnership

Abstract

A group of consecutive 4 genes from the microorganism *Kytococcus sedentarius* (Ksed_09150 – Ksed_09190) 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, and cellular localization data. 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 database.

Introduction

Kytococcus sedentarius is a Gram positive bacterium in the genus Kytococcus. It is an opportunistic pathogen. The genome has been sequenced and contains 2,785,024 bp, with a G+C content of 71.6%. This genome encodes 2639 protein coding genes. This organism is classified as a chemoheterotroph, as it requires methionine and several other amino acids for growth. It is strictly aerobic.

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 monens in 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).

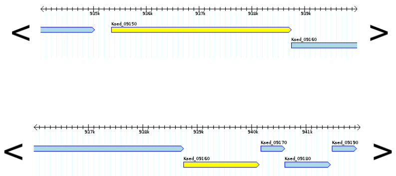

Figure 1. The loci of the containing genes annotatated in this project.

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?

Results

Ksed_09150:

Initially GENI-ACT predicted that the gene produced Isoleucyl-tRNA Synthetase. This was supported by the top BLAST hits for the amino acid sequence. Following after entering the protein sequence and concluding that it is located in the cytoplasm because there was no evidence of transmembrane helices in TMHMM or Phobious and no signal peptide was detected in signal P. Phobious and Psortb predict with 9.97/10 probability that it was in the cytoplasm. As such the proposed annotation is Isoleucyl-tRNA Synthetase. The reaction catalyzed is illustrated in the figure above.

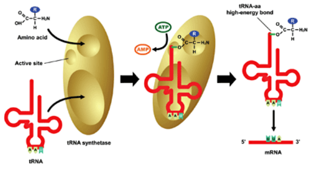

Figure 2. Illustration of how Isoleucyl-tRNA Synthetase catalyze the attachment of isoleucine to its transfer RNA.

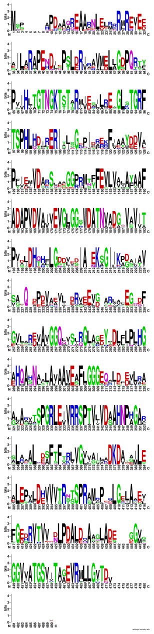

Figure 3. The WebLogo obtained from Ksed_09160.

Kytococcus sedentarius 09160:

This gene was initially proposed by GENI-ACT as Dihydrofolate synthase. This gene can be located within the cytoplasm, with the Cytoplasmic Membrane score (obtained from the PSORT-B system) was greater than 7. After researching this gene, its function is to catalyze the following reation:

ATP + dihydropterate + L-glutamate = ADP + phosphate + dihydrofolate.

This gene's WebLogo design shows that there are plenty of hydrophobic amino acids (shown in black in Figure 3) found within the protein.

Ksed_091090:

The initial proposed product of this gene by GENI-ACT was nucleoside diphosphate kinase. 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, and the cytoplasmic location. The cytoplasmic localization of this enzyme was predicted confirmed by a PSORT score of 9.97. Nucleoside diphosphate kinases are enzymes that catalyze the exchange of terminal phosphate between different nucleoside diphosphates as shown in figure 4 below.

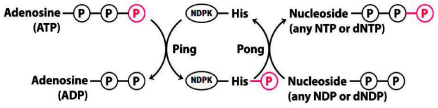

Figure 4. An example reaction schematic for a nucleoside diphosphate kinase.

Ksed_09180:

The initial proposed product of this gene by GENI-ACT was Undecaprenyl-diphosphatase. 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 results from TMHMM predicts six trans-membrane helices the results are shown tin the figure above. PSORT predicts that the protein is localized in the cytoplasmic membrane (score 10.0). This enzyme is involved in the biogenesis of bacterial cell-wall polysaccharides, such as peptidoglycan, cytoplasmic synthesized

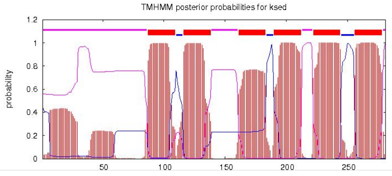

Figure 5. TMHMM output for Ksed_09180 showing the presence of 6 transmembrane helixes.

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	Gene-Act Gene Products	Proposed Annotation
09150	Isoleucyl- tRNA Synthetase	Isoleucyl- tRNA Synthetase
09160	Dihydrofolate Synthase	Dihydrofolate Synthase
09180	Undecaprenyl-diphosphatase 2	Undecaprenyl-diphosphatase 2
09190	Nucleoside Diphosphate Kinase	Nucleoside Diphosphate Kinase

References

Manat et al. (2014) Deciphering the Metabolism of Undecaprenyl-Phosphate: The Bacterial Cell-Wall Unit Carrier at the Membrane Microbial, Drug Resistance Volume 20, 119-214

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

Acknowledgments

Supported by NSF ITEST Strategies Award Number 1311902