Annotation of the *Kytococcus sedentarius* Genome from Locus Tags Ksed_08590 to Ksed_08610

Emma Pryll*, Jonathan McDaneld* and Gabriella Rivera* and Cheryl Smith

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

Abstract
A group of three consecutive genes from the microorganism Kytococcus sedentarius (Ksed_08590, Ksed_08600, Ksed_08610) were annotated using the NCBI genome annotation website GENI-ACT. The Genbank proposed gene product names for each gene that were assed in terms of the basic genomic information, sequence-based similarity data, structure-based evidence, cellular localization data, and alternative open reading frame. The Genbank proposed gene product name and coordinates did not differ significantly from the information found for Ksed_08600, Ksed_08610 and as such, the genes appear to be correctly identified in the database. The Genbank proposed gene coordinates did differ significantly from the gene coordinates determined for Ksed_08590.

Introduction
*Kytococcus sedentarius* is a free-living, gram-positive bacteria, isolated from a marine environment. According to ZoBell and Upham (1944), it’s known for its role as an opportunistic pathogen causing diseases like valve endocarditis and hemorrhagic pneumonia. *Kytococcus sedentarius* is an aerobic, spherical organism. It sometimes appears in tetrads and may be surrounded by a slimy layer on smear. *Kytococcus sedentarius* is the only known producer of the antibiotics monensin A and B, that have been isolated from varying environments, including human skin and groundwater.

In terms of taxonomy, *Kytococcus sedentarius* is in the kingdom *Bacteria* and it’s phylum and class are both *Actinobacteria*. *K. sedentarius* is of interest because it is a part of the not very well studied family, *Dermacoccaceae*. Though the genes of *Kytococcus sedentarius* may have been studied before, it has not been manually annotated until now.

Modules of the GENI-ACT (http://www.geni-act.org) were used to complete *Kytococcus sedentarius* genome annotation. The modules are described below:

### Modules

<table>
<thead>
<tr>
<th>Module</th>
<th>Activities</th>
<th>Questions Investigated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Module 1 - Basic Information</td>
<td>DNA Coordinates and Sequence, Protein signatures</td>
<td>Are the genome coordinates and sequence information correct?</td>
</tr>
<tr>
<td>Module 2 - Sequence-Based Similarity Search</td>
<td>Basic, CDD, T-Coffee, Weblogo</td>
<td>Is the sequence similar to other sequences in GenBank?</td>
</tr>
<tr>
<td>Module 3 - Cellular Localization Data</td>
<td>Gram Stain, TMHMM, SignalP, PSORT, Photoscan</td>
<td>Is my sequence similar to other sequences in GenBank?</td>
</tr>
<tr>
<td>Module 4 - Alternative Open Reading Frame</td>
<td>I-MUG Sequence Viewer For Alternative ORF Search</td>
<td>Has the amino acid sequence of my protein gene been correctly predicted by the computer?</td>
</tr>
<tr>
<td>Module 5 - Structure-Based Evidence</td>
<td>T-Blasfem, Pfam, PSG</td>
<td>Are there functional domains in my protein?</td>
</tr>
</tbody>
</table>

### Results

*Kytococcus sedentarius* 08590:

The initial proposed product of this gene by GENI-ACT was DNA binding response regulator. This proposed gene product was supported by the top BLAST hits. The gene appeared to be very well conserved when compared to orthologs such as *Mycobacterium* as evidenced by WebLogo particularly between residues 232-233. The cellular localization indicated the gene product is cytoplasmic. The predicted start codon location of the gene was 873078 but lacked a Shine-Dalgarno sequence. There was another possible location for a start codon indicated in Figure II. As predicted, the gene product is DNA binding response regulator.

*Kytococcus sedentarius* 08600:

The initial proposed product of this gene by GENI-ACT was histidine kinase. Histidine kinase is a multifunctional protein in the transerase class of enzymes that play a role in signal transduction across the cellular membrane. 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, cellular localization indicated two helices across the cell membrane in Phobius and TMHMM. HMM Logo showed the gene was well conserved. As such, the proposed gene product is histidine kinase.

*Kytococcus sedentarius* 08610:

The initial proposed product of this gene was a Lipo LqBP beta propeller. The BLAST results indicated a possible hypothetical protein. A hypothetical protein is a protein that has been predicted but for which there is no experimental evidence. TMHMM predicted no transmembrane helices while Phobius and SignalP predicted a non-cytoplasmic signal peptide. The proposed gene product for Ksed_08610 is a Lipo LqBP beta propeller.

### 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 the genes appear to be correctly annotated by the computer database.

### Acknowledgments

Supported by NSF I/TEST Strategies Award Number 1311902