**Annotation of the *Kytococcus sedentarius* Genome from Locus Tags Ksed_07570 to Ksed_07590**

Ehsan Rafique, Fatimah Salman, Hannah Schumacher, and Omer Bitar

Lackawanna High School, and the Western New York Genetics in Research Partnership

---

**Abstract**

The investigation of three genes (Ksed_07570, Ksed_07580, and Ksed_07560) from the marine bacteria, *Kytococcus sedentarius*, have been manually annotated with the assistance of the GENI-ACT (Genomics Educational National Initiative – Annotation Collaboration Toolkit). The Genbank database was an automated method for predicting gene presence and function. Manual annotation is used to confirm, reject, or reassign the name and function of the investigated gene. Annotation was done by co-mapping proteins to their organisms by examining similarity in amino acid sequences (sequence based similarity), structure based similarity, and cellular localization data. In most cases, the annotated genes did not differ from the Genbank predicted function.

**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:

<table>
<thead>
<tr>
<th>Module</th>
<th>Description</th>
<th>Activities</th>
<th>Questions Investigated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Module 1: Basic Information</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Module 2: Sequence-Based Signal peptide Data</td>
<td>Blast, T-coffee, WebLog</td>
<td>Is my sequence similar to other sequences in Genbank?</td>
<td></td>
</tr>
<tr>
<td>Module 2: Cellular Localization Data</td>
<td>Gene Scan, T-coffee, SignalP, Phobius</td>
<td>Is my protein a signal peptide, secreted or embedded in the membrane?</td>
<td></td>
</tr>
<tr>
<td>Module 3: Alternative Open Reading Frame</td>
<td>Rieg Sequence Viewer For Alternative ORF Search</td>
<td>Has the amino acid sequence of my protein been called correctly by the computer?</td>
<td></td>
</tr>
</tbody>
</table>

**Results**

*Kytococcus sedentarius* (Ksed_07570)

The initial proposed product of this gene by GENI-ACT was a molecular chaperone, GroEL. 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 proposed annotation is a RNA-binding protein.

**Conclusion**

Except for Ksed_07450, 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.

**References**

Sims et al. (2009). Complete genome sequence of *Kytococcus sedentarius* type strain (5411T). Standards in Genomic Sciences, 12, 20.

**Acknowledgments**

Many thanks to Dr. Stephen Koury, Dr. Rama Dey-Rao, and Danielle Wilson for their guidance, support and dedication to this project. Supported by NSF REU Strategies Award Number 1311962