Using the Annotation of the *K. sedentarius* Genome to enhance STEM Experiences

Christopher Lozada, Tyler Allen, Zachary VanKoevering, and Shawn Flanagan

Biology And Advanced Science Explorations, Newark High School, Newark, NY and the Western New York Genetics in Research Partnership

Abstract

A group of consecutive 7 genes from the microorganism *Kytococcus sedentarius* (Ksed_04170 – Ksed_04230) 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, the possibility of horizontal gene transfer, and the production of an RNA product. 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 r database.

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>Modules</th>
<th>Activities</th>
<th>Questions Investigated</th>
<th>Modules</th>
<th>Activities</th>
<th>Questions Investigated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Module 1: Basic Information Module</td>
<td>Codon Coordinates and Sequence, Protein Sequence, Impact is the sequence of my gene and protein? Where is it located in the genome?</td>
<td>Module 2: Sequence-Based Similarity Data</td>
<td>Blast, CDD, T-Coffee, Weblogo</td>
<td>Is my sequence similar to other sequences in GenBank?</td>
<td></td>
</tr>
<tr>
<td>Module 3: Cellular Localization Data</td>
<td>Gram Stain, TMHMM, SignalP, PHFPhilos</td>
<td>Module 4: Alternative Open Reading Frame</td>
<td>MG Sequence Viewer For Alternate ORF Search</td>
<td>Has the amino acid sequence of my protein been called correctly by the computer?</td>
<td></td>
</tr>
<tr>
<td>Module 5: Structural-Based Evidence</td>
<td>TIGRfam, Pfam, PDB</td>
<td>Module 6: Enzymatic Function</td>
<td>KEGG, MetaFun, E.C. Numbers</td>
<td>Are there functional domains in my protein?</td>
<td></td>
</tr>
<tr>
<td>Module 7: Gene Duplication/ Gene Degeneration</td>
<td>Peptidoglycan, Peptidylglycine</td>
<td>Module 8: Evidence for Horizontal Gene Transfer</td>
<td>Phylogenetic Tree</td>
<td>In what process does my protein take part?</td>
<td></td>
</tr>
<tr>
<td>Module 9: RNA</td>
<td>RPAM</td>
<td>Module 10: Advanced Genomic Annotation</td>
<td></td>
<td>Are there other forms of my gene in the bacterium?</td>
<td></td>
</tr>
</tbody>
</table>

### Results

Throughout the project presented to us by UB my group would be varied in our motivation to complete the project. But within our procrastination was an underlying urge to learn more. One reason we felt a little bit behind is because we would ask so many offshoot questions and get sidetracked on the science of the cells of *Kytococcus sedentarius*. We wanted to know everything we could about trans-membranous proteins and a chemoheterotroph was used as an organism of interest for this project. This bacterium was selected because it has yet to have been thoroughly studied utilizing bioinformatics (Sims et al., 2009).

During the project, two activities piqued the group’s interest. WebLogo and Transmembrane topology graphing were of interest.

**Figure 1A:** a WebLogo from Ksed_04530

**Figure 1B:** a WebLogo from Ksed_04700

**Figure 2:** trans-membrane topology maps from the 3 different genes.

### Conclusion

By having after-school STEM activities available to students, it increases awareness for the types of jobs that are out there involving STEM. For science programs in high schools can be slow and drawn-out because of the variety in the learning speeds of students. The setting of the classroom requires the instructor to go at one pace for everyone, but in after-school activities instructors can better individualize how instruction is happening. Within my group, 2 of us were 3 modules ahead of the rest of the group, but in the end we all learned what was necessary. This way of instruction gives more flexibility to how someone is educated. Personally I feel as if I have learned more from after-school STEM activities than I have in a classroom setting for the same area.

### References


### Acknowledgments

Supported by NSF ITEST Strategies Award Number 1311902

The participants gratefully thank Stephen Koury, Rama Dey-Rao, Patricia Masso-Welch, Jennifer Cooke, the Rural Area Health Education Center, and the University at Buffalo for their assistance and guidance.