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

Maddison Knapp and Lon Knappenberger
Westfield Academy and Central School, 203 East Main Street, Westfield, New York 14787 and the Western New York Genetics in Research Partnership

Abstract

*Kytococcus sedentarius* is a Gram-positive coccus. A specific gene was annotated from this bacteria to find its function. The database Geni-Act was used to determine what the gene was. Then the data that was found from the gene details page of Geni-Act was used in several different websites. Eight Geni-Act modules were used to gather this information. The results show a protein found in the cytoplasmic. With the help of Geni-Act and several other resources, it was determined that the gene product was a ubiquione biosynthesis.

Introduction

*Kytococcus sedentarius* is a Gram-positive bacterium that is arranged in packets of eight in a cubical shape (1). *Kytococcus sedentarius* was found on a partially submerged swing set in San Diego California. It is a coccolid or spheroidal shaped bacteria that can live on the skin (1).

The genome of *Kytococcus sedentarius* was 541,785,024 bp long and codes for 2,703 genes. Of the 2,703 genes, 2,638 were protein-coding genes. 72% of these genes were assigned with a predicted function while the remaining one were annotated as hypothetical proteins.

The gene that we investigated, Ksed_02230 had the DNA coordinates of 201238, 221912 containing 675 base pairs. The section codes for 224 amino acids.

The purpose was to look and verify information from GENI-ACT, and to better understand a poorly studied family within the tree of life. The other purpose was to determine a possible function of the gene product and provide an opportunity for original research.

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>
</tr>
</thead>
<tbody>
<tr>
<td>Module 1: Basic Information</td>
<td>DNA Coordinates and Frames, Protein Sequence</td>
<td>What is the sequence of my gene and protein? Where is it located in the genome?</td>
</tr>
<tr>
<td>Module 2: Sequence-Based Similarity Data</td>
<td>Basic, CID, T-Coffee, WebLogo</td>
<td>Is my sequence similar to other sequences in Geni-ACT?</td>
</tr>
<tr>
<td>Module 3: Cellular Localization Data</td>
<td>Gram Stain, TMHMM, SigPred, Phobius</td>
<td>Is my protein in the cytoplasm, secreted or embedded in the membrane?</td>
</tr>
<tr>
<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>
</tr>
<tr>
<td>Module 5: Structure-Based Evidence</td>
<td>T-Raptor, Plant, PSB</td>
<td>Are there functional domains in my protein?</td>
</tr>
<tr>
<td>Module 6: Enzymatic Function</td>
<td>KEGG, MetaCyc, KEGG</td>
<td>Is the product of my gene active?</td>
</tr>
<tr>
<td>Module 7: Gene Duplication</td>
<td>Protein, Pseudogene</td>
<td>Are there other forms of my gene in the bacterium? Is my gene functional?</td>
</tr>
<tr>
<td>Module 8: Evidence for Homologous Gene Transfer</td>
<td>Phylogenetic Tree</td>
<td>Has my gene co-evolved with other genes in the genome?</td>
</tr>
</tbody>
</table>

Results

**Figure 1.** This is the top protein BLAST hit for the non-redundant database. Which shows a ubiquione biosynthesis protein found in *Kytococcus* (evalue = 8e-66).

**Figure 2.** This is the second protein BLAST hit. It is also a ubiquione biosynthesis protein but found in *Sarminococcus chungangensis* (evalue = 1e-66).

**Figure 3.** This is the three-dimensional figure of Protein Data Base top hit, SAM-dependent methyltransferase from *Corynebacterium glutamicum* Methanol methyltransferases are involved in the synthesis of coenzyme Q or ubiquione.1

**Figure 4.** This is the result for the Psorts B for Ksed_02230. This shows the location of the gene product in the cytoplasmic.

**Figure 5.** Phylogenetic tree shows the *Sarminococcus* as the closest neighbor to *Kytococcus*. Both *Sarminococcus* and *Kytococcus* are classified in the same family. This indicates that Ksed_02230 did not arise in *Kytococcus* by horizontal gene transfer.

Conclusion

The GENI-ACT proposed gene product for Ksed_02230 is ubiquione one biosynthetic which is found in the cytoplasmic. Ubiquione (UQ) is also known as a coenzyme, which is a redox-active lipophilic (combining with or dissolving) molecule that is found in all cells. Also it is in the membranes of many organelles where it participates in a variety of processes. In the inner membrane of the mitochondria, UQ diffuses freely and is necessary for the function of the electron transport chain (ETC), as it enables the transfer of electrons from mitochondrial complexes I and II and mitochondrial complex III.2

References


Acknowledgments

This project was supported by a National Science Foundation Innovative Technology Experience for Students and Teachers (TEST) Award (#011902).