Annotation of the *Kytococcus sedentarius* Genome from Ksed _08920 to Ksed _08950

Joshua Ninan, Erin Washburn*, Madison Crowley* and Meghan Janora

Cattaraugus Little Valley School and the Western New York Genetics in Research Partnership

### Abstract

Four genes from the microorganism *Kytococcus sedentarius* (Ksed _08920 – Ksed _09950) were annotated using the collaborative genome annotation website GENI-ACT. The Genbank proposed gene product name for two of the genes was assessed using the proposed genome modules. Each gene product was assessed using a differential system in terms of their 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 and enzymatic function. Thus far, the GenBank-proposed gene product names did not differ significantly from the proposed gene annotation for the two annotated genes in the group, and as such, the genes appear to be correctly annotated by the i-database for the modules that were completed on each.

### Introduction

*Kytococcus sedentarius*, the only known producer of the antibiotic monearin A and B, has been isolated from varying environments, including human skin, groundwater, and even airplane cabins. It can be a human opportunistic pathogen. Strain Ksed_09474, the type strain, is a free-living, nonmotile, Gram-positive bacterium, originally isolated from a marine environment in about 1944. It grows as spherical to oval and occurs predominantly in tetrad which can be arranged in cubic packets. It is non-ensapulated and does not form endospores, thus being aerobic and chemotrophotrophic, requiring methionine and a tetramethylvaleric acid for growth and grows well in McIlvan nutrient broth. To 10% (w/v). According to Sims et al. (2009), *Kytococcus sedentarius* is a microorganism of interest for several reasons. This bacterium is the natural source of the antibiotics monearin A and monearin B (Sims et al., 2009). *Kytococcus sedentarius* has been implemented as the biological agent of a number of opportunistic infections including staphylococci, histoplasmosis, and dental infections (Sims et al., 2009). Finally, the phylogeny of this microorganism is of interest, as it is a member of the family Dermatococaceae within the actinobacterial suborder Micrococcaceae, which has yet to have been thoroughly studied utilizing bioinformatics (Sims et al., 2009).

### Methods and Materials

Modules of the GENI-ACT were used to complete *Kytococcus sedentarius* genome annotation.

### Results

**Basic Information**

- **ID**: Ksed
- **GenBank**: MT286047
- **Length**: 527
- **Amino acid length**: 322

**Sequence Similarity-Based Evidence**

- **COGs**: COG3514 (lipid transport and metabolism)
- **TIGR**: The strain does not grow in a cell-free medium.
- **REBDO**: Actin and cytoskeletal proteins.

**Structure Based Evidence**

- **TOPMOL**: 1H2A 1994
- **Pfam**: 1994
- **HMMER**: HMM profiles for Pfam, COG, KEGG, and MAP
- **BLAST**: Predicted protein function is consistent with the Pfam

**Cellular Location Data**

- **Transporter Family**: Transporters
- **Signal Peptide**: No signal peptide predicted
- **SwissProt**: Predicted protein function is consistent with the Pfam

**Alternative Open Reading Frame**

- **EGG**: NA

### Conclusion

The initial proposed product of this gene by GENI-ACT was a acyl-CoA thioesterase. This gene product proposal was supported by the top BLAST hit for the amino acid sequence, acyl-CoA thioesterase II. Sequence (Ksed_08920) similarity supports the proposal by what is shown in the semi-conserved portion of the WebLogos. Not shown in Figure 3 are similarly conserved regions at the beginning of the sequence as well. The COF file used in this study was the TIGRFAF database as reviewed in TIGR00189 (GS_08940). The second gene (Ksed_08940) is a CoA thioesterase which can hydrolyze a broad range of acyl-CoA thioesters. Its physiological function may not be known. Figure 3 shows the PSORTB data to predict where the protein is soluble in the cytoplasm and integral membrane protein, or secreted by the cell. The score of 7.50 predicts it in the cytoplasm. This prediction is also supported by the fact that the proposed product is an acyl-CoA thioesterase with possible function in lipid transport and metabolism.

**Figure 3** – The WebLogos N321-N384 for *Kytococcus sedentarius* 08920, where most conservation is shown. Other conserved portions not shown were at A49, N76, and N77.

### References

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

### Acknowledgments

Supported by NSF I/SECST Strategies Award Number 1311902

Figure 1. Assaying electron micrograph of *Kytococcus sedentarius*

Figure 2. Location of the locus tags analyzed in this project

Figure 3 – The WebLogos N321-N384 for *Kytococcus sedentarius* 08940, where most conservation is shown. Other conserved portions not shown were at A49, N76, and N77.

Figure 4 – *Kytococcus sedentarius* 08940 Protein results showing the score of 7.5 as cytoplasmic.

Figure 5 – Crystal structure of acyl-CoA thioesterase from Yersinia pestis in complex with acyl-CoA A. The proposed product of gene 08940 of *Kytococcus sedentarius* may have similar structure.