Gene Annotation of Ksed_04940, a Putative Copper Chaperone Protein in Kytococcus sedentarius

Lucas Haskin and Lon Knappenberger

Westfield Academy and Central School, Westfield, New York 14787 and The Western New York Genetics in Research Partnership

Abstract

Kytococcus sedentarius is a Gram-positive coccici. A gene with the locus tag of Ksed_04940 in this bacteria's genome was annotated to verify computer called information and look for possible function for the gene product. Eight GENI-Act modules were used to gather this information. The results show the Ksed_04940 protein is likely to be extracellular, but it does not have a signal peptide. The top BLAST hits of Ksed_04940 all are copper binding proteins, whereas the top COG hit is for copper chaperone CypD. These findings support the computer call gene of Ksed_04940 as a copper binding protein. Further the COG hit suggests that Ksed_04940 is a copper binding involved in copper metal ions into intracellular compartments or to copper dependent enzymes.

Introduction

Kytococcus sedentarius is a Gram positive bacterium that is arranged in packets of eight in a cubical shape. Kytococcus sedentarius was found on a partially submerged swing set in San Diego California. It is a coccoid or spherical shaped bacterium that can live on the skin. The genome of Kytococcus sedentarius is 541,725,024 bp long and codes for 2,703 genes. Of the 2,703 genes predicted, 2,639 were protein-coding genes. 72% of these genes were assigned with a predicted function while the remaining ones were annotated as hypothetical proteins.

Materials and Methods

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

- **Module 1: Basic Information**
  - **DNA Coordinates and Sequence:** Used to determine the DNA coordinates of the gene.
  - **What is the sequence of my protein?**

- **Module 2: Sequence-Based Similarity Data**
  - **BLAST:** The BLAST tool was used to find the best hit of the gene.
  - **SignalP:** Predicts the signal peptide.

- **Module 3: Cellular Localization Data**
  - **Gene Targets:** Predicts the subcellular localization of the gene.
  - **Is my protein the cytoplasm, secreted or embedded in the membrane?**

- **Module 4: Alternative Open Reading Frame**
  - **MG Sequence Viewer For Alternative ORF Search:** Helps determine if the gene has multiple reading frames.

- **Module 5: Structure-Based Evidence**
  - **TOPFam, Pfam, PROD:** Are these functional domains in my protein?

- **Module 6: Enzymatic Function**
  - **KEGG, MetaCyc, E.C. Number:** In what process does my protein take part?

- **Module 7: Gene Duplication / Gene Degeneration**
  - **Paralog, Pseudogene:** Are there other forms of my gene in the bacterium? Is my gene functional?

- **Module 8: Evidence for Horizontal Gene Transfer**
  - **Phylogenetic Tree:** Have my gene and other genes in the genome have been linked to human diseases such as Menkes disease.

- **Module 9: RNA**
  - **RFAM:** Does my gene encode a functional RNA?

Results

Figure 1. A scanning electron micrograph of Kytococcus sedentarius growing in the characteristic cuboidal packet form (Sims et al., 2009).

Figure 2. The results of a protein BLAST using the Non-redundant database shows a Multi-species copper-binding protein as the top hit (e-value 2e-65).

Figure 3. SignalP results predict whether Ksed_04940 has a signal peptide. No signal peptide is predicted.

Figure 4. An image of the three-dimensional structure of the top hit of Ksed_04940 in the Protein Data Base. This is a solution structure of a putative copper-ion-binding protein from Bacillus anthracis.

Figure 5. This is the WebLogo of the T-Coffee multisequence alignment for Ksed_04940. It shows that our gene is highly conserved through most of the sequence.

Figure 6. The Pfam-B section of the gene notebook, and this shows the location of the gene product as extracellular. This makes sense based on the purpose of chaperone proteins.

Figure 7. PFAM pairwise alignment shows a heavy metal association domain is conserved. (e = value = 8.1e-13)

Figure 8. This is the phylogenetic tree for the... It shows Micrococcus as the closest neighbor. Kytococcus and Micrococcus are the same down to taxonomic family. This indicates Heredity and Lateral gene transfer.

Figure 9. This is the ortholog neighborhood viewer of the gene Ksed_04930. It shows evidence of transfer by heredity. The neighborhood is conserved in Kytococcus and its common ancestors Micrococcus and Pseudonocardia.

Conclusions

The top BLAST hit is a Copper-binding protein, which is extracellular. Copper-binding proteins chaperone copper ions to copper-dependent enzymes and protect the cytoplasm during the transport of metal ions. Mutations to these proteins in Humans have been linked to human diseases such as Menkes disease (growth failure) and Wilson's disease (liver disease).

Figure 10: Illustration of a copper insertion chaperone and transporter component from Bacillus subtilis

Literature Cited


Acknowledgements

This project was supported by a National Science Foundation Innovative Technology Experiences for Students and Teachers (ITEST) Award (#1311902).

Image 1: A scanning electron micrograph of Kytococcus sedentarius growing in the characteristic cuboidal packet form (Sims et al., 2009).

Image 2: The results of a protein BLAST using the Non-redundant database shows a Multi-species copper-binding protein as the top hit (e-value 2e-65).

Image 3: SignalP results predict whether Ksed_04940 has a signal peptide. No signal peptide is predicted.

Image 4: An image of the three-dimensional structure of the top hit of Ksed_04940 in the Protein Data Base. This is a solution structure of a putative copper-ion-binding protein from Bacillus anthracis.

Image 5: This is the WebLogo of the T-Coffee multisequence alignment for Ksed_04940. It shows that our gene is highly conserved through most of the sequence.

Image 6: The Pfam-B section of the gene notebook, and this shows the location of the gene product as extracellular. This makes sense based on the purpose of chaperone proteins.

Image 7: PFAM pairwise alignment shows a heavy metal association domain is conserved. (e = value = 8.1e-13)

Image 8: This is the phylogenetic tree for the... It shows Micrococcus as the closest neighbor. Kytococcus and Micrococcus are the same down to taxonomic family. This indicates Heredity and Lateral gene transfer.

Image 9: This is the ortholog neighborhood viewer of the gene Ksed_04930. It shows evidence of transfer by heredity. The neighborhood is conserved in Kytococcus and its common ancestors Micrococcus and Pseudonocardia.

Image 10: Illustration of a copper insertion chaperone and transporter component from Bacillus subtilis.