Annotation of the *Kytococcus sedentarius* Genome of Locus Tags Ksed_14030, Ksed_20690, and Ksed_15070

Abdulelah Ahmed, Kevin Grimes, Fadel Saleh, and Omer Bitar
Lackawanna High School, Lackawanna, NY and The Western New York Genetics in Research Partnership

Abstract

The investigation of three genes (Ksed_14030, Ksed_20690, and Ksed_15070) 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 uses an automatic method of predicting gene presence and function. Manual annotation is used to confirm, reject, or reassign the name and function of the investigated gene. Annotation is achieved by comparing proteins of other organisms by examining similarities in amino acid sequences (sequence based similarity data), structure based evidence, and cellular localization data. In most cases, the annotated genes did not differ from the Genbank predicted function.

Introduction

Kytococcus sedentarius is a strictly aerobic, non-motile, non-encapsulated, and non-endospore forming gram positive coccoid bacterium, found predominantly in tetrad formation. This organism is classified as a chemoheterotroph, as it requires methionine and several other amino acids for growth. Originally isolated from a microscope slide submerged in sea water in 1944, *Kytococcus sedentarius* grows well in sodium chloride at concentrations less than 10% (w/v).

According to Sims et al. (2009), *Kytococcus sedentarius* is a microorganism of interest for several reasons. This bacterium is a natural source of the oligoketide antibiotics monensin A and monensin B (Sims et al., 2009). *Kytococcus sedentarius* has been implemented as the etiological agent of a number of opportunistic infections including valve endocarditis, hemorrhagic pneumonia, and pitted keratolysis (Sims et al., 2009). Finally, the phylogeny of this microorganism is a source of interest, as it is a member of the family *Dermacoccaceae* within the actinobacterial suborder *Micrococcineae*, which has yet to have been thoroughly studied utilizing bioinformatics (Sims et al., 2009).
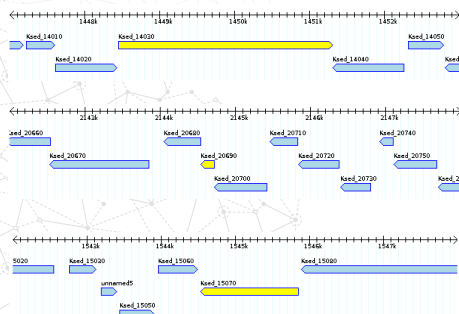

Figure 1. The locus tags and relative position of the genes under investigation in this research

Methods

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

Modules	Activities	Questions Investigated
Module 1- Basic Information Module	DNA Coordinates and Sequence, Protein Sequence	What is the sequence of my gene and protein? Where is it located in the genome?
Module 2- Sequence-Based Similarity Data	Blast, CDD, T-Coffee, WebLogo	Is my sequence similar to other sequences in Genbank?
Module 3- Cellular Localization Data	Gram Stain, TMHMM, SignalP, PSORT, Phobius	Is my protein in the cytoplasm, secreted or embedded in the membrane?
Module 4- Alternative Open Reading Frame	IMG Sequence Viewer For Alternate ORF Search	Has the amino acid sequence of my protein been called correctly by the computer?
Module 5- Structure-Based Evidence	TIGRfam, Pfam, PDB	Are there 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 Degradation	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,	Has my gene co-evolved with other genes in the genome?
Module 9- RNA	RFAM	Does my gene encode a functional RNA?

Results

Ksed_14030:

The initial proposed product of this gene by GENI-ACT was a RNA helicase. This gene product proposal was supported by the top BLAST hits for the amino acid sequence, the presence of well-curated protein functional domains within the amino acid sequence, the transmembrane topography of the amino acid sequence, and the cellular location of the amino acid sequence. As such, the proposed annotation is a RNA helicase.
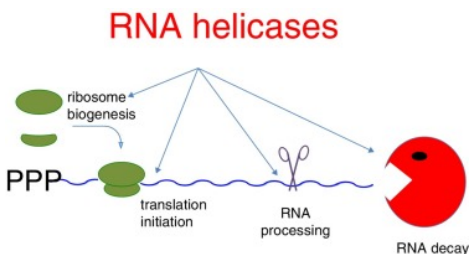

Figure 2 – the various roles of RNA helicases (sciencedirect.com)

Ksed_20690:

The initial proposed product of this gene by GENI-ACT was a 50s ribosomal protein L30. 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 50s ribosomal protein L30.
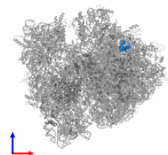

Figure 3 – 3-D model of a 50s ribosomal protein L30 in E. coli (ebi.ac.uk)
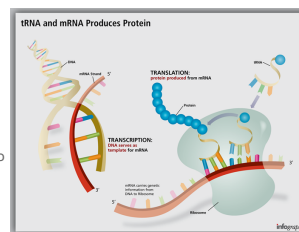

Figure 4 – Ribosomal proteins are involved in the process of translation, when RNA is converted in to Amino Acids (infographics.com)
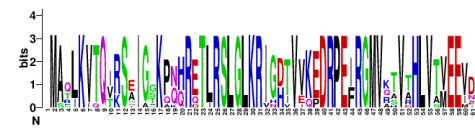

Figure 5 – the entire Ksed_20690 WebLogo results fit on one line, showing highly conserved amino acids

Ksed_15070:

The initial proposed product of this gene by GENI-ACT was aminopeptidase N. 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 aminopeptidase N
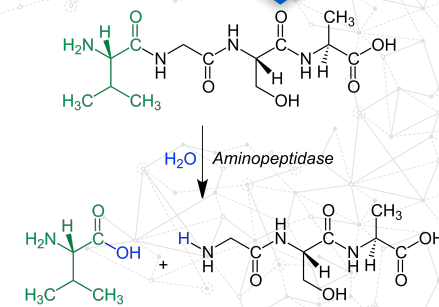

Figure 6 – aminopeptidases remove the peptide bond between amino acids from proteins. (wikimedia.org)

Conclusion

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.

Gene Locus	Geni-Act Gene Products	Proposed Annotation
Ksed_14030	RNA Helicase	RNA Helicase
Ksed_20690	50s ribosomal protein L30	50s ribosomal protein L30
Ksed_15070	Aminopeptidase N	Aminopeptidase N

References

Sims et al. (2009). Complete genome sequence of *Kytococcus sedentarius* type strain (541T). Standards Genomic Sciences,12 - 20.

Roles of RNA helicase (sciencedirect.com)

3-D model of a 50s ribosomal protein L30 in E. coli (ebi.ac.uk)

RNA translation (infographics.com)

Aminopeptidase reaction (wikimedia.org)

Acknowledgments