University at Buffalo *The State University of New York* | REACHING OTHERS

Gene Annotation of Ksed_2300, a Putative Protein Tyrosine Phosphatase(PTP) in *Kytococcus sendentarius*

Ally Kassman 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 cocci. A section of the bacteria's genome was annotated to verify information and look for possible functions for the gene product. Eight GENI-ACT modules were used to gather this information. The results show a protein found in the cytoplasm. This is thought to be signal transduction protein tyrosine phosphatase. They play a very important role in cellular signaling

Introduction

Kytococcus sedentarius is a Gram positive bacterium that is arranged in packets if eight in a cubical shape [4]. Kytococcus sedentarius was found on a partially submerged swing set in San Diego California [5]. It is coccoid or a spherical shaped bacterium that can live on the skin [4].

The genome of *Kytococcus sedentarius* 541 2,785,024 bp long and codes for 2,703 genes. Of the 2,703 genes predicted, 2,639 were protein-coding genes. 72.1% of these genes were assigned with a predicted function while the remaining ones were annotated as hypothetical proteins [6].

The gene that we investigated, Ksed_2300 had the DNA coordinates of 227593..228096 containing 504 base pairs. This section codes for 167 amino acids.

The purpose of this was to help learn more about *Kytococcus sedentarius* because it is not a well known type of bacteria. The information from GENI-ACT help us in understanding the gene better.

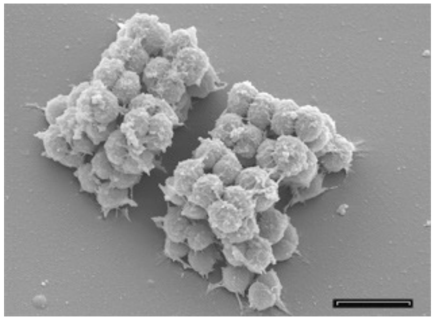


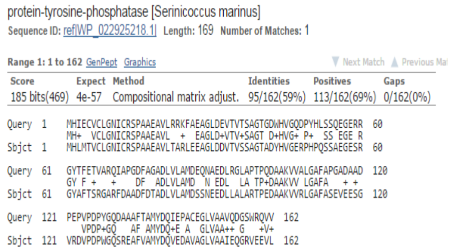
Figure 1: This is an image of Kytococcus Sedentarius
Source: Standards in Genomic Science (2009) **1**:12-20

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:

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?

Results



protein-tyrosine-phosphatase [Serinicoccus marinus]
Sequence ID: ref|WP_022925218.1| Length: 169 Number of Matches: 1

Range 1: 1 to 162 GenPept Graphics

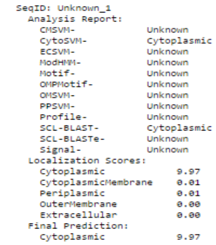
Score	Expect	Method	Identities	Positives	Gaps
185 bits(469)	4e-57	Compositional matrix adjust.	95/162(59%)	113/162(69%)	0/162(0%)

```
Query  1    MHIECVCLGNICRSPAAEAVLRRKFAEAGLDEVTVTSAGTGDWHVGQDPYHLSSQEGERR  60
            MH+  VCLGNICRSPAAEAVL   +  EAGLD+VTV+SAGT D+HVG+ P+   SS EGE R
Sbjct  1    MHLMTVCLGNICRSPAAEAVLTARLEEAGLDDVTVSSAGTADYHVGERPHPQSSAEGESR  60

Query  61   GYTFETVARQIAPGDFAGADLVLAMDEQNAEDLRGLAPTPQDAAKVVALGAFAPGADAAD  120
            GY F +    +    DF  ADLVLAMD  N EDL  LA TP+DAAKVV LGAFA   + +
Sbjct  61   GYAFTSRGARFDAADFDTADLVLAMDSSNEEDLLALARTPEDAAKVVRLGAFASEVEESG  120

Query  121  PEPVPDPYGQDAAAFTAMYDQIEPACEGLVAAVQDGSWRQVV  162
               VPDP+GQ   AF AMYDQ+E A  GLVAA++ G   +V+
Sbjct  121  VRDVPDPWGQSREAFVAMYDQVEDAVAGLVAAIEQGRVEEVL  162
```

Figure 2: This is the best hit in UniProtKB/SwissProt database, (Protein Tyrosine Phosphatase, e-value 4e-57)



```
SeqID: Unknown_1
  Analysis Report:
    CMSVM-              Unknown
    CytoSVM-            Cytoplasmic
    ECSVM-              Unknown
    ModHMM-             Unknown
    Motif-              Unknown
    OMPMotif-           Unknown
    OMSVM-              Unknown
    PPSVM-              Unknown
    Profile-            Unknown
    SCL-BLAST-          Cytoplasmic
    SCL-BLASTe-         Unknown
    Signal-             Unknown
  Localization Scores:
    Cytoplasmic                 9.97
    CytoplasmicMembrane         0.01
    Periplasmic                 0.01
    OuterMembrane               0.00
    Extracellular               0.00
  Final Prediction:
    Cytoplasmic                 9.97
```

Figure 3: This is the Psort B section of the gene notebook and it shows that the location of the gene product is in the Cytoplasm.

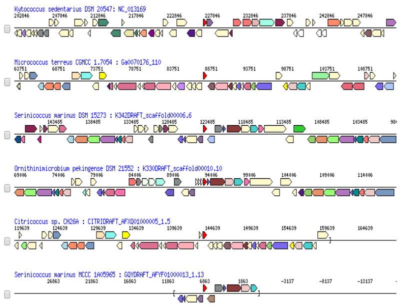


Figure 4: This is the ortholog neighborhood viewer of the gene Ksed_2300. It shows evidence that Ksed_2300 (in red) was compared to orthologs. In this case, Ksed-2300 is not flanked by similar genes

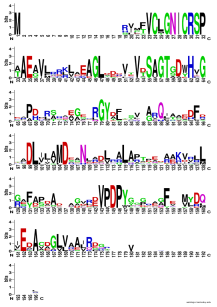


Figure 5: This graph is the Web Logo for the T-COFFEE of the orthologs of Ksed_2300. It shows that there are not very many matches because most are medium size letters instead of large.

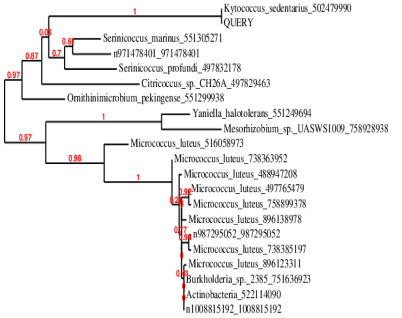


Figure 6: This is the Phylogenetic Tree Ksed_2300 and it shows that *Kytococcus* and *Serinicoccus* have the same order but a different family. Unlike Figure 4, this shows heredity as a source of the gene

Figure 7: This is an image from the Protein database of what Ksed_2300 could potentially look like. It is a Phosphotyrosine phosphatase from *Vibrio cholera* (e-value of 1.47273E-16).

Conclusion

Protein tyrosine phosphorylation is the putative gene product, the protein is a fundamental regulatory mechanism controlling cell proliferation, differentiation, communication, and adhesion. Disruption of this key regulatory mechanism contributes to a variety of human diseases including cancer, diabetes, and auto-immune diseases.[1]

Protein Tyrosine Phosphatases (PTP's) represent a large family of enzymes. They play a very important role in cellular signaling within and between cells. PTP's work with Protein Tyrosine Kinases (PTKs) to regulate signal transduction in a cell. PTKs phosphorylate tyrosine residues on a substrate protein and PTPs remove these phosphates from substrate tyrosines (dephosphorylation). Since the phosphorylation status of a protein can modulate its function, PTKs and PTPs work together to regulate protein function in response to a variety of signals, including hormones, mitogens, and oncogenes.[2]

Protein tyrosine phosphatases (PTPs) are a group of enzymes that catalyze the removal of phosphate groups from tyrosine residues by the hydrolysis of phosphoric acid monoesters. They directly oppose the actions of kinase and phosphorylate enzymes.[3]

References

1
http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3421019/
2
http://earth.callutheran.edu/Academic_programs/Departments/BioDev/omm/ptp1b/molmast.htm
3
https://www.tocris.com/pharmacologicalBrowswer.php?ItemId=187767#.Vx9o7z9WI_s
4
http://www.vumicro.com/vumie/help/VUMICRO/Kytococcus_Sedentarious.htm
5 http:/biocyc.org/KSED478801/organism-Fummary?object=KFED478801
6 http://standardsingenomics.org

Acknowledgments

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