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

Gene Annotation of Ksed_02310, a Putative Fructosamine-3-Kinase in *Kytococcus sedentarius*

Taylor R. Kassman and Lon Knappenberger

Westfield Academy and Central School, Westfield NY, 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 function for the gene product. Eight GENI-ACT modules were used to gather this information. The putative gene product is a fructosamine kinase. The protein is found in the cytoplasm and contains no transmembrane helices. The protein adds a phosphoryl group ($-PO_3^{2-}$) to a fructosamine (a compound between a sugar and an amine, which contains nitrogen).

Introduction

Kytoccus 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 [2]. It is a coccoid or spherical shaped bacterium that can live on the skin [1].

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.

The gene that we investigated, Ksed_02310, had the DNA coordinates of 228093..228911 containing 819 base pairs. The section codesfor 272 amino acids.

The purpose was to look for and verify information from GENI-ACT. This helps us to better understand a poorly studied family within the bacterial suborder *Micrococcineae*, and to determine a possible function of the gene product and provide an opportunity for original research.

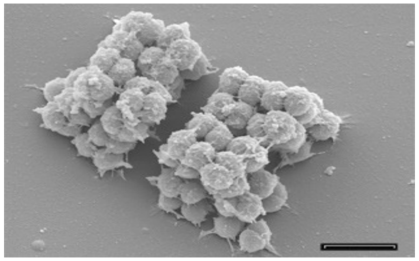


Figure 1. Kytococcus Sedentarius
Source: Standards in Genomic Science (2009) **1**:12-20

Methods and Materials

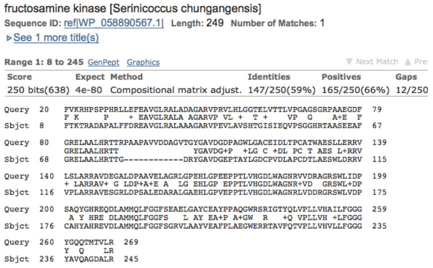
Modules of the GENI-ACT (http://www.geni-act.org/) were used to complete *Kytococcus sedentarius* genome annotation . The modules are described below:

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

Kytococcus sedentarius 02310:

The initial proposed product of this gene by GENI-ACT was a fructosamine kinase . This gene product proposal was supported by the top BLAST hits for the amino acid sequence using both the non redundant and Swiss prot data bases. The top COG hit was a Fructosamine-3-kinase. The top TIGRfam hit was a kinase. The top Pfam was a fructosamine kinase PF03881. As shown in the figure below, the top PDB data base hit was a fructosamine-3-kinase.



fructosamine kinase [Serinicoccus chungangensis]
Sequence ID: ref|WP_058890567.1| Length: 249 Number of Matches: 1
▸See 1 more title(s)

Range 1: 8 to 245 GenPept Graphics

Score	Expect	Method	Identities	Positives	Gaps
250 bits(638)	4e-80	Compositional matrix adjust.	147/250(59%)	165/250(66%)	12/250

```
Query  20   FVKRHPSPPHRLLEFEAVGLRALADAGARVPRVLHLGGTELVTTLVPGAGSGRPAAEGDF  79
            F K      P     + EAVGLRALA AGARVP VL +   T +     VP  G    A+E  F
Sbjct  8    FTKTRADAPALFFDREAVGLRALAAAGARVPEVLAVSHTGISIEQVPSGGHRTAASEEAF  67

Query  80   GRELAALHRTTRPAAPAVVDDAGVTGYGAVDGDPAGWLGACEIDLTPCATWAESLLERRV  139
            GRELAALHRTT                 YGAVDG+P  +LG C +DL PC T AES L+RRV
Sbjct  68   GRELAALHRTTG------------DRYGAVDGEPTAYLGDCPVDLAPCDTLAESWLDRRV  115

Query  140  LSLARRAVDEGALDPAAVELAGRLGPEHLGPEEPPTLVHGDLWAGNRVVDRAGRSWLIDP  199
            + LARRAV+ G LDP+A+E A  LG EHLGP EPPTLVHGDLWAGNR+VD  GRSWL+DP
Sbjct  116  VPLARRAVESGRLDPSALEDARALGAEHLGPVEPPTLVHGDLWAGNRLVDDRGRSWLVDP  175

Query  200  SAQYGHREQDLAMMQLFGGFSEAELGAYCEAYPPAQGWRSRIGTYQLVPLLVHAILFGGG  259
             A Y HRE DLAMMQLFGGFS   L AY EA+P A+GW  R   +Q VPLLVH +LFGGG
Sbjct  176  CAHYAHREVDLAMMQLFGGFSGRVLAAYVEAFPLAEGWERRTAVFQTVPLLVHVLLFGGG  235

Query  260  YGQQTMTVLR  269
            Y  Q   LR
Sbjct  236  YAVQAGDALR  245
```

Figure 2. Alignment for top BLAST hit on NR database. E-value 4e-80.

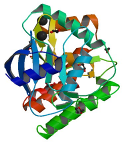


Figure 3: Protein Data Base hit for a fructoamine -3-kinase found in *Thermobifida Fusca*. E- Value = 8.09825E-26

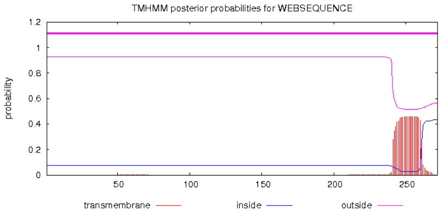


Figure 4- TMHMM graphical output of Ksed_02310.

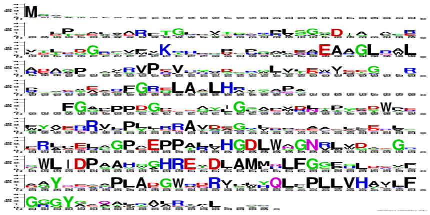


Figure 5- This graph is the Web Logo for the T COFFEE alignment of orthologs of Ksed_02310. It shows that the gene is conserved.

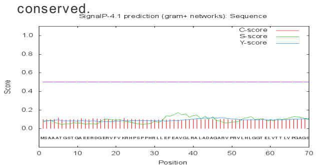


Figure 6- This graph represents the Signal Peptide. Signal Peptide probability is .105. No cleavage sites.

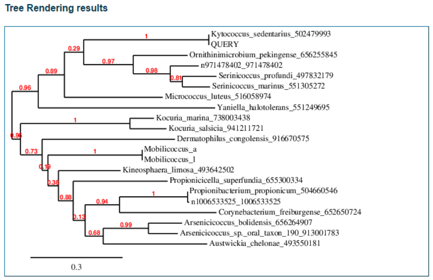


Figure 7- Phlogenetic tree for Ksed_02310. It shows the *Ornithimicrobium pekingense* as the closest neighbor.

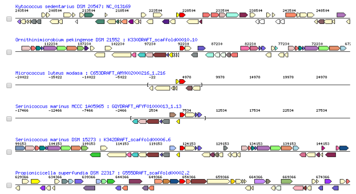


Figure 8- This is the ortholog neighborhood viewer of the gene Ksed_02310. It shows evidence of transfer by heredity. The neighborhood is conserved in close relatives of *Kytococcus*.

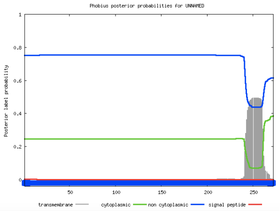


Figure 9- This graph represents the Phobius output. Number of predicted transmembrane helices is zero. No signal peptide. Cytoplasmic score of 7.50 so my gene product is in the cytoplasm.

Conclusions

In molecular biology, the fructosamine kinase family is a family of enzymes. This family includes eukaryotic fructosamine-3-kinase enzymes which may initiate a process leading to the deglycation of 1-deoxy-1-morpholinofructose, fructoselysine, fructoseglycine, fructose and glycated lysozyme. The family also includes ketosamine-3-kinases (KT3K). Ketosamines derive from a non-enzymatic reaction between a sugar and a protein. Ketosamine-3-kinases (KT3K) catalyse the phosphorylation of the ketosamine moiety of glycated proteins. The instability of phosphorylated ketosamine leads to its degradation, and KT3K is thus thought to be involved in protein repair (2).

The function of the prokaryotic members of this group has not been established. However, several lines of evidence indicate that they may function as fructosamine-3-kinases (FN3K). First, they are similar to characterized FN3K from mouse and human. Second, the Escherichia coil members are found in close proximity on the genome to fructose-6-phosphate kinase (PfkB). Last, FN3K activity has been found in the blue-green algae *Anacystis montana* indicating such activity-directly demonstrated in eukaryotes-is nonetheless not confined to eukaryotes (2).

References

1. http://www.vumicro.com/vumie/help/VUMICRO/ Kytococcus_Sedentarious.htm.
2. http://biocyc.org/KSED478801/organismFumma ry?o8

Acknowledgments

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