

# Gene Annotation of Ksed\_04930, a Putative Signal Transduction Histidine Kinase in *Kytococcus sedentarius*

David Grayson 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 function for the gene product. Eight GENI-ACT modules were used to gather this information. The results show a protein found in the cytoplasmic membrane with two transmembrane helices. This gene product is thought to be signal transduction histidine kinase involved in adding a phosphate group to a protein.

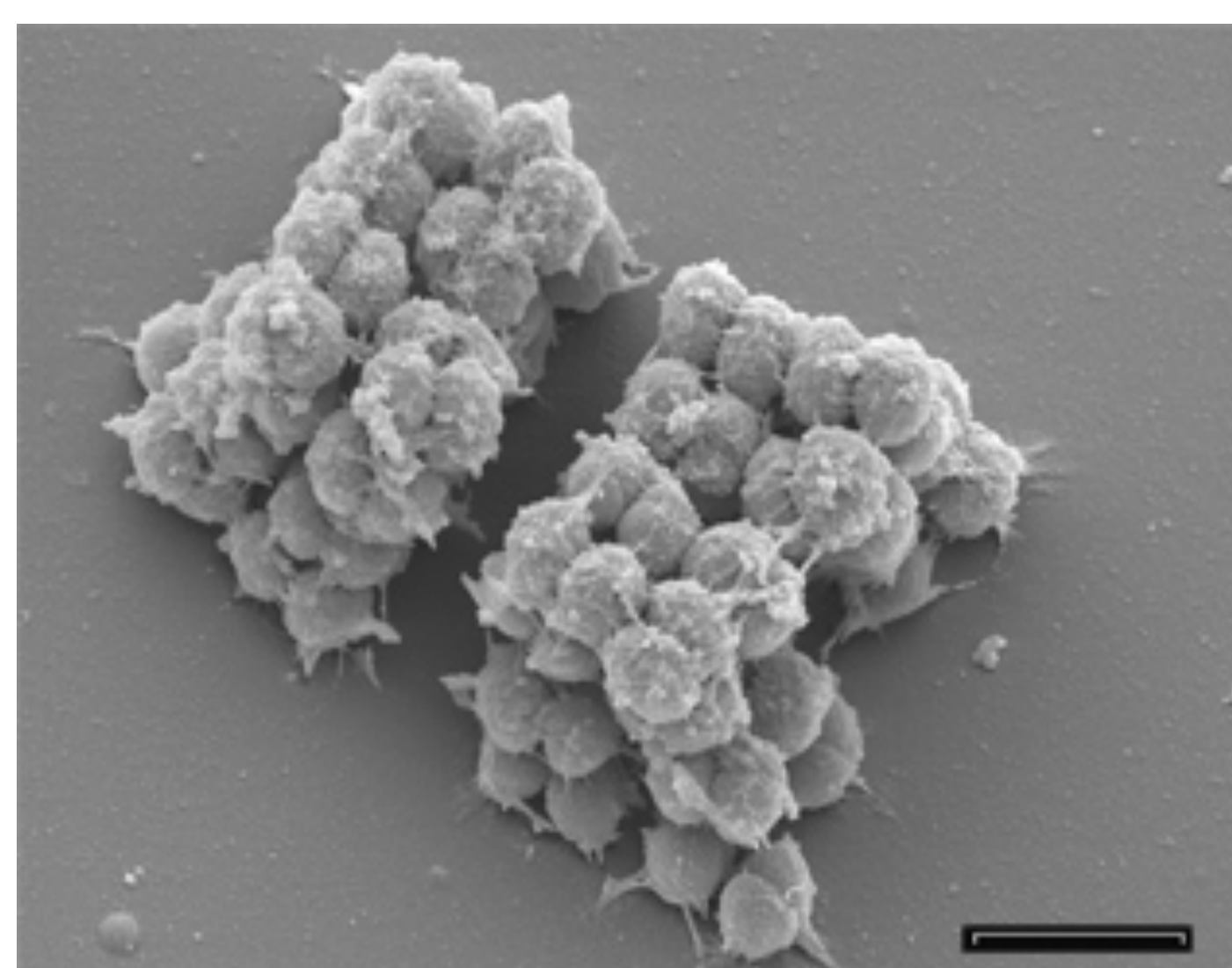
## Introduction

*Kytococcus 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 coccid 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\_04930, had the DNA coordinates of 488333.489502 containing 1170 base pairs. The section codes for 389 amino acids.

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



*Kytococcus sedentarius*

## 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:

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

RecName: Full=Signal transduction histidine-protein kinase BaeS [Escherichia coli K-12]  
 Sequence ID: sp|P30847.2|BAES\_ECOLI Length: 467 Number of Matches: 1

Range 1: 187 to 461 GenPept Graphics

Score	Expect	Method	Identities	Positives	Gaps
175 bits(444)	3e-48	Compositional matrix adjust.	97/276(35%)	163/276(59%)	2/276(0%)

Query 92 RRIQSPINALQAVKMGAGRYDTRVPAMGAGTVEEDLAGFFNTMAARLQRTEDTRNML 151  
 R + \*P+ L + \*A G + ERV + E+ LA FH +A+ L++ + RR +  
 Sbjct 187 RGLLAPVKRLVDGTHKLAAGDFTTRVTP-SEDELKLAQDFHQLASTLEKQQRDRFM 245

Query 152 SDLAHLRTPVTVYHEALQDQVSHDEPPELGMGLARLRLVEDINQVRAEQR 211  
 +D++HELCT++VL EAQQV + \* ++ LA-LVLD++ +S ++E2  
 Sbjct 246 ADISHLARTPLAVRGELEALQDQVRFPTVAVLQAEVTLKLVLDLQLMSDEGA 305

Query 212 IELDPSPOVSGLLHAATEAKKAYATGVALTMADPGIAYVDVDRMGQVIGLNLTHA 271  
 + \*P+ + LL A A +E +A+G+ L + V DR R+ Q+ NLL N+  
 Sbjct 306 LAYQAPVDLIPLELVAGGAFRRFASRGLKIQSLPDSITVFGDRDLMLFNNLENS 365

Query 272 LRHTPAGQRTLEARGSPGVLEVDVTDGDISPELPHVFERFRTGDFARDRDGQGV 331  
 LA+T +G2 + + A Q V LA D+ Q+S + L +FERFER +\*R+R GGG+  
 Sbjct 366 LRVDSSGSLQISAGQDKVRLTFADSNAPVSDQKQLKLFERFRTGDSRNRASGSG 425

Query 332 GLAISRALIQAGG-TLSATSTTAGATFTDILPTQ 366  
 GLAI +++AH G ++A S G + T++LP ++  
 Sbjct 426 GLAICLNIVEARNGRIIAHSPFGVSIIVLPLER 461

Figure 1. Top BLAST hit in UniProtKB/SwissProt database (histidine protein kinase, (e-value 3-48).

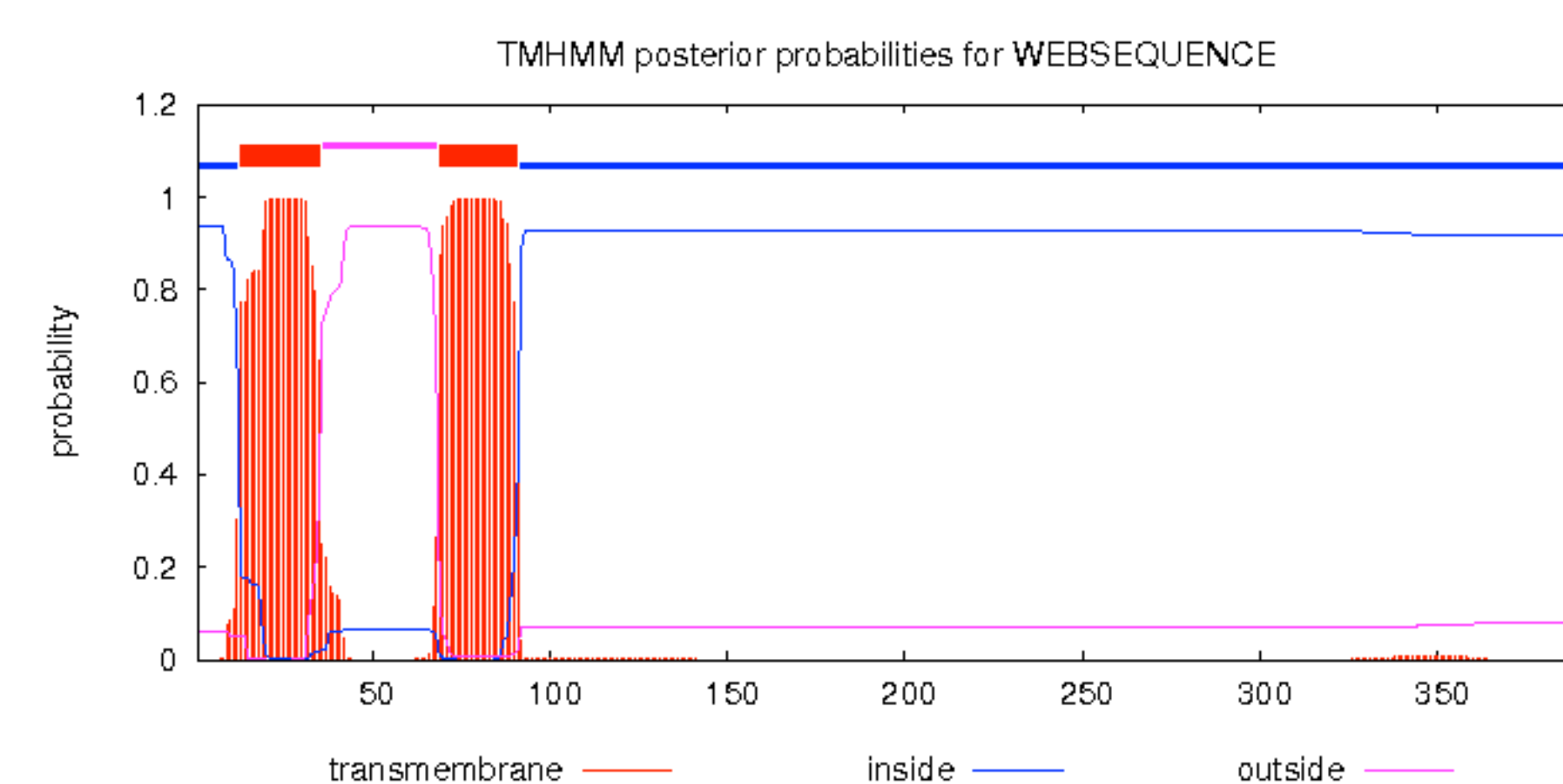


Figure 2. TMHMM graph predicting 2 transmembrane helices in Ksed\_04930.

## Results

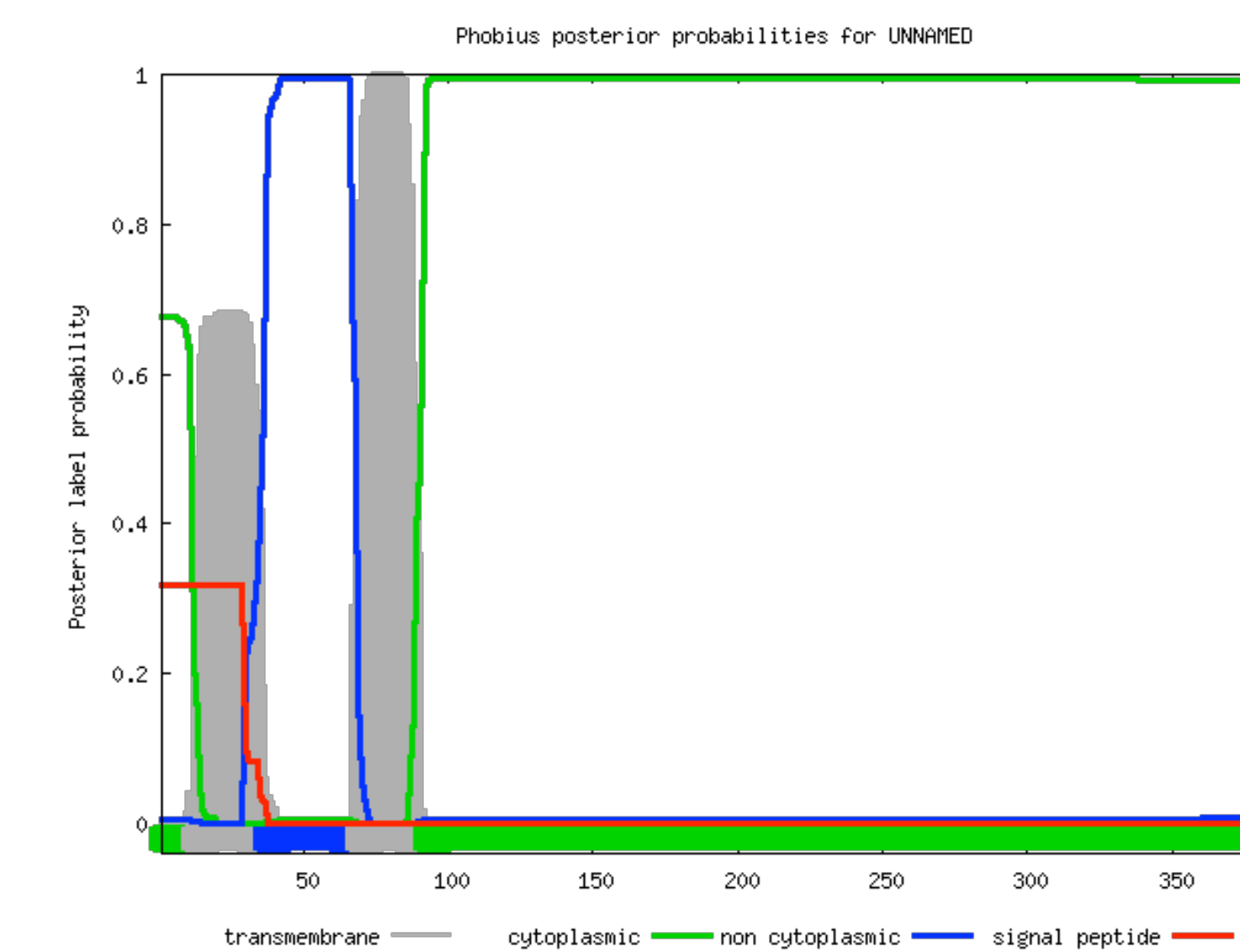


Figure 2. This graph represents the Phobius output. The graph confirms the transmembrane helices found on the THMM site.

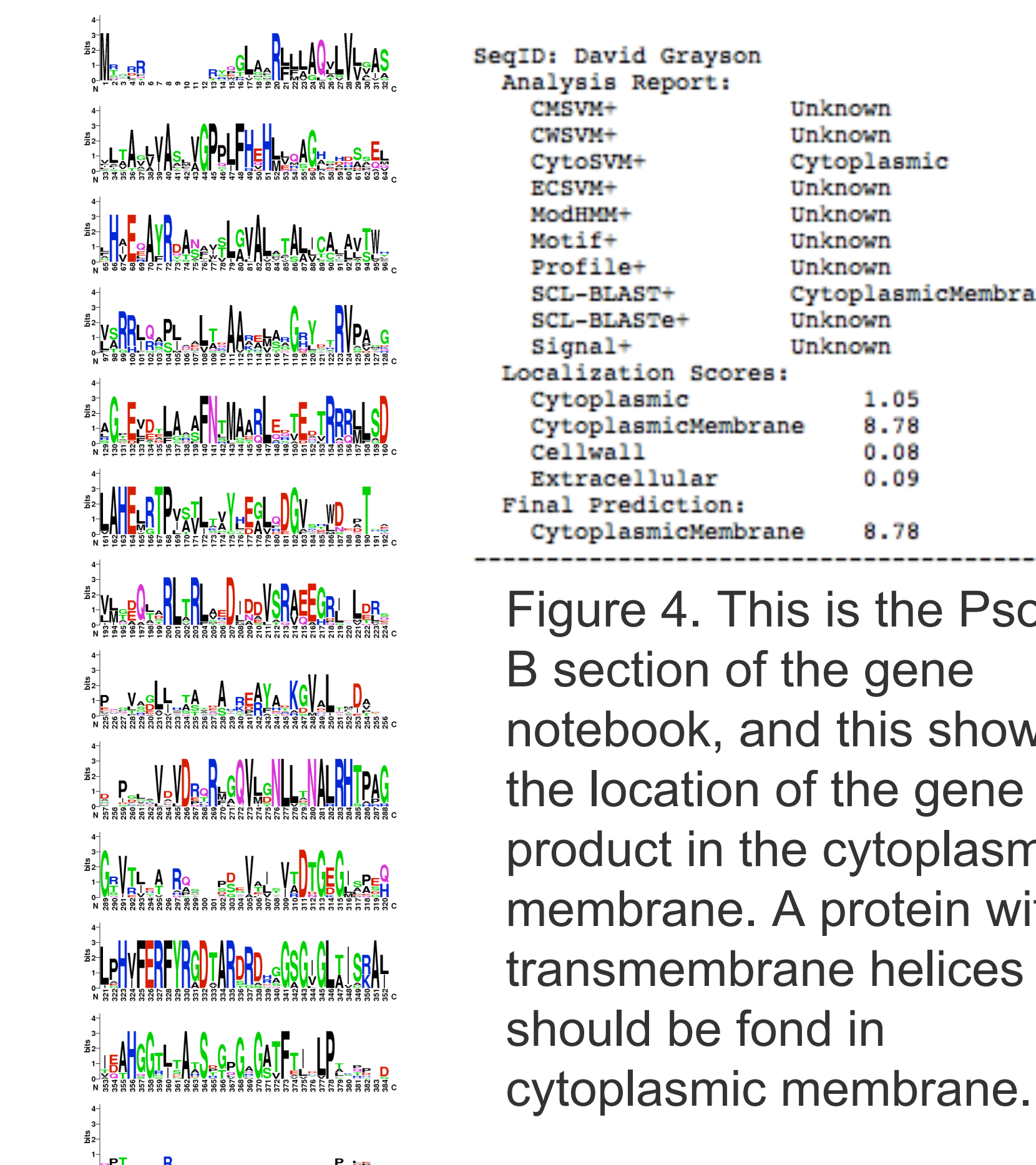


Figure 3. This graph is the Web logo for the T-COFFEE of the orthologs of Ksed\_04930 . It shows that the gene is conserved except at the C-terminus.

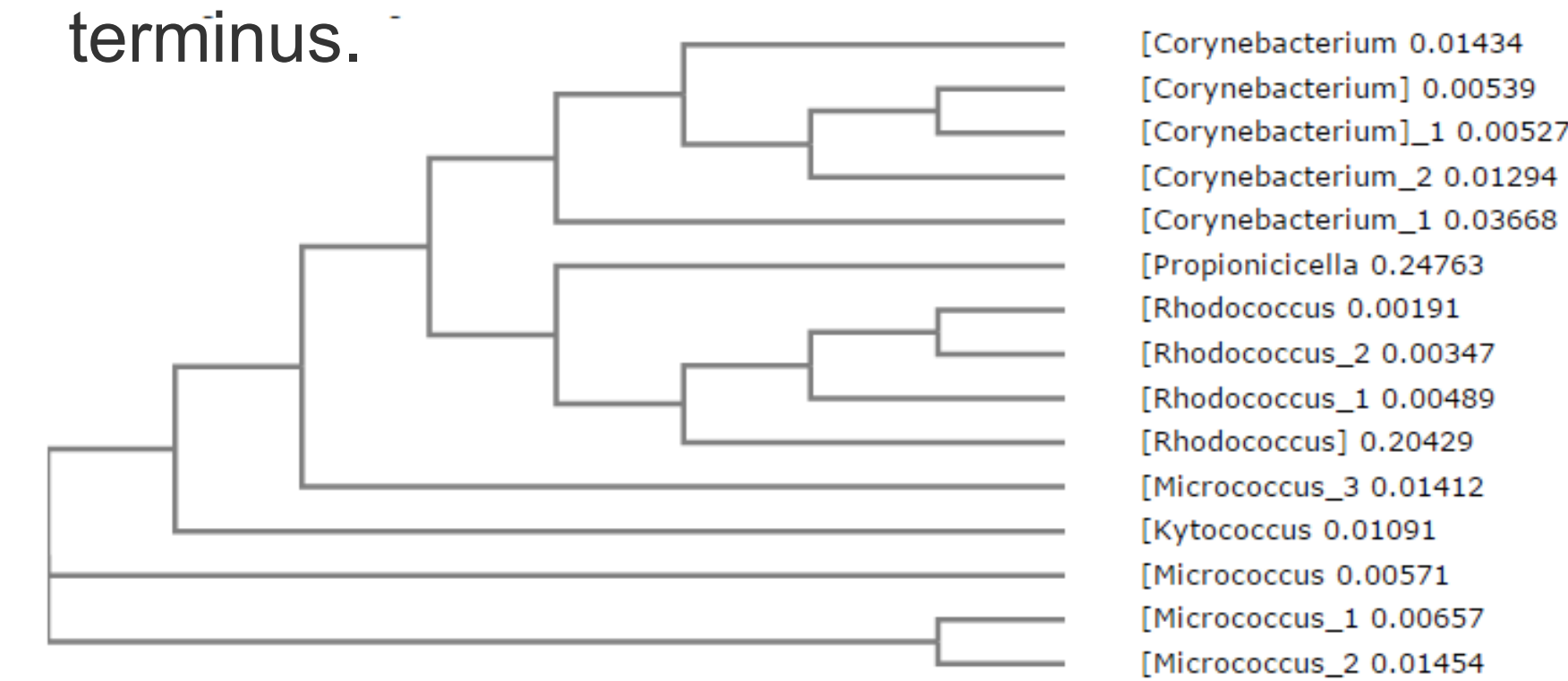


Figure 5. Phylogenetic tree for Ksed\_04930 . It shows *Micrococcus* as the closest neighbor. *Kytococcus* and *Micrococcus* are the classified the same to the family level. This indicates Ksed\_04930 likely did not arise in *Kytococcus* by lateral gene transfer.

## Results continued

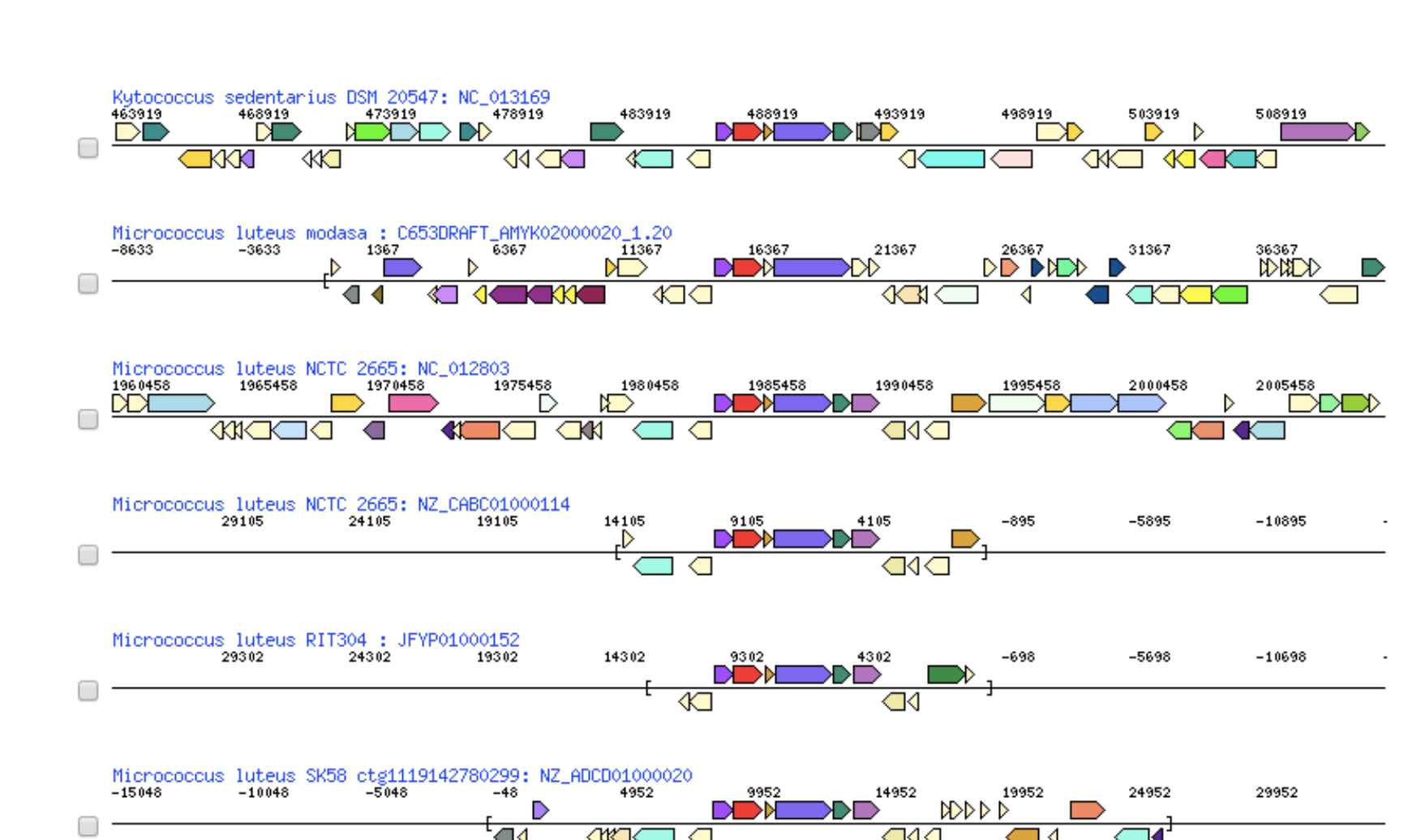


Figure 6. This is the ortholog neighborhood viewer of the gene Ksed\_04930. It shows evidence of transfer by heredity. The neighborhood is conserved in close relatives of *Kytococcus*.

## Conclusions

The Gene product for ksed\_04930 appears to be a Histidine protein kinase and its main purpose is to catalyze the reaction bonding a phosphate group to a protein.

ATP + protein L-histidine becomes ADP + protein N-phospho-L-histidine.

The two substrates of this enzyme are ATP and protein L-histidine, and the two products are ADP and protein N-phospho-L-histidine ([http://en.wikipedia.org/wiki/Histidine\\_kinase](http://en.wikipedia.org/wiki/Histidine_kinase)). These proteins are used to transmit signals and control complex processes within the cells.

## Literature Cited

- [http://www.vumicro.com/vumie/help/VUMICRO/Kytococcus\\_Sedentarius.htm](http://www.vumicro.com/vumie/help/VUMICRO/Kytococcus_Sedentarius.htm)
- <http://biocyc.org/KSED478801/organism-Fummary?object=KFED478801>

## Acknowledgements

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