

Annotation of the *Kytococcus sedentarius* Genome at Locus Tags

Ksed_RS05035 Ksed_RS05045 and Ksed_RS05050

Ryan Beckett, Jennie Merciliott, Jessica Perdue, Abdulrahman Yassein, and Kate Johnston
Niagara Falls High School, Niagara Falls New York and The Western New York Genetics in Research and Health Care Partnership

Abstract

Three genes from the microorganism *Kytococcus sedentarius* (Ksed_RS05035, Ksed_RS05045, and Ksed_RS05050) were annotated using the collaborative genome annotation website GENI-ACT. The Genbank proposed gene product name for each gene was assessed using its basic genomic information, amino acid sequence-based similarity data, structure-based evidence from the amino acid sequence, cellular localization data, potential alternative open reading frames, enzymatic function, presence or absence of gene duplication and degradation, the possibility of horizontal gene transfer. The original proposed gene product name did not differ significantly from the proposed gene annotation for each of the genes in the group and therefore the genes appear to be correctly annotated in the database.

Introduction

Kytococcus Sedentarius is an aerobic gram positive bacteria originally isolated from a marine environment (Sims et al., 2009). Its cells are spherical and in the form of tetrads that are groups in cubic packets (Sims et al., 2009). The optimal growth temperature is between 23-36°C (Sims et al., 2009). It is a bacterium of interest due to the diseases and conditions it has been linked to playing a role in causing. It is considered the cause of skin infections that degrade human calluses, endocarditis and hemorrhagic pneumonia (Sims et al., 2009).

Kytococcus Sedentarius can produce oligoketide antibiotics monensins A and B, and *Kytococcus* appear to have significant survival abilities when airborne (Folayan, A. et al., 2018). While the bacteria can cause life threatening illness, it is not clear how Kytococci enter the body, and they are typically resistant to penicillin G, methicillin, and isoaxazol penicillins (Folayan, A. et al., 2018). More studies are needed to help prevent and treat diseases caused by *Kytococcus Sedentarius*.

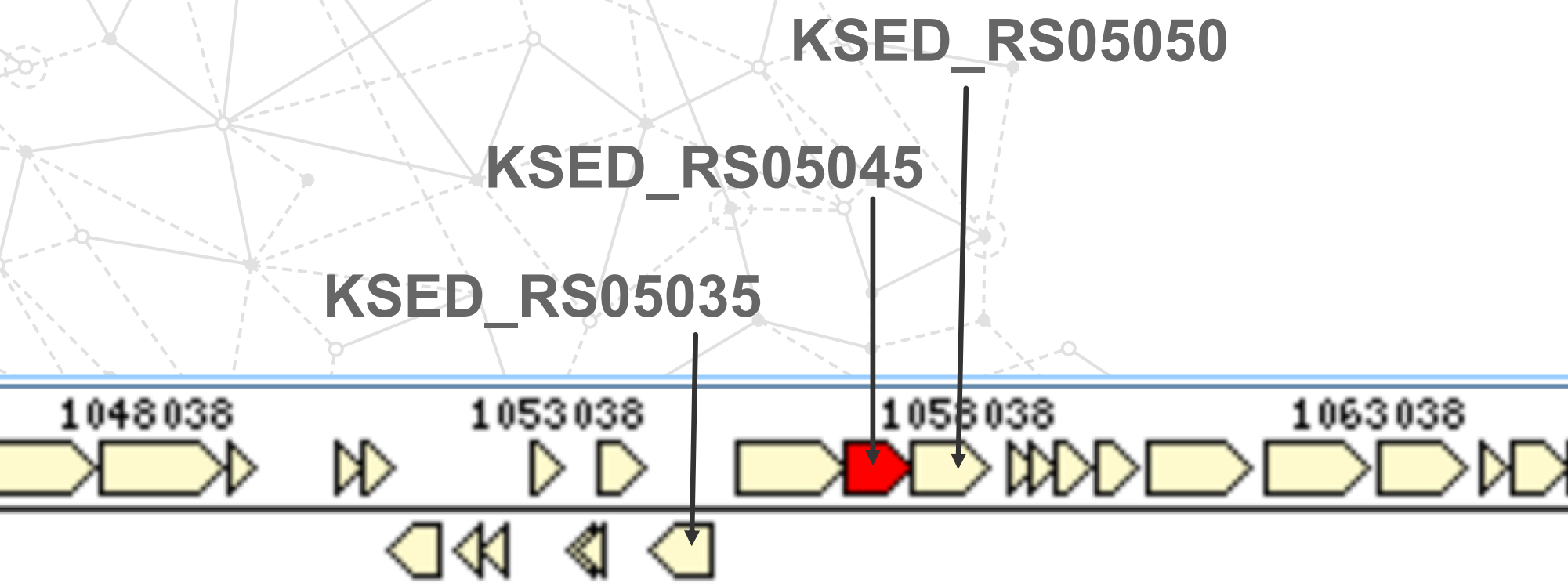


Figure 1 - The locus tags and relative position of the genes under investigation in this research

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_RS05035:

The original gene predicted by GENI-ACT was a Beta-Hydroxybutyrate Dehydrogenase. This gene product proposal was supported by the top BLAST hits for the amino acid sequence, and the presence of well-curated protein functional domains within the amino acid sequence. Its function appears to be to synthesize or degrade ketone bodies. Based on this annotation, the originally predicted gene of Beta-Hydroxybutyrate Dehydrogenase correlates with these findings.

Ksed_RS05045:

The initial proposed product of this gene by GENI-ACT was a type II secretion system protein F. This gene product proposal was consistent with the top BLAST hits for the amino acid sequence. The T-coffee showed alignment with the sequence. TMHMM and Phobius indicate that it functions as a transmembrane protein. And so, the proposed annotation is a type II secretion system protein F.

Ksed_RS05050:

The initial proposed product of these genes by GENI-ACT was Flp pilus assembly protein TadB. This gene product matches the results from the BLAST sequence. The sequence obtained similar results between the T-Coffee and Web Logo sequences. This protein showed evidence of being a transmembrane protein. Every module completed supported the initial proposal that the gene is a Flp pilus assembly protein.

Figure 2 – Ksed_RS05035 PDB result produced this 3-D model of the protein D-beta-hydroxybutyrate dehydrogenase from *Sinorhizobium meliloti*

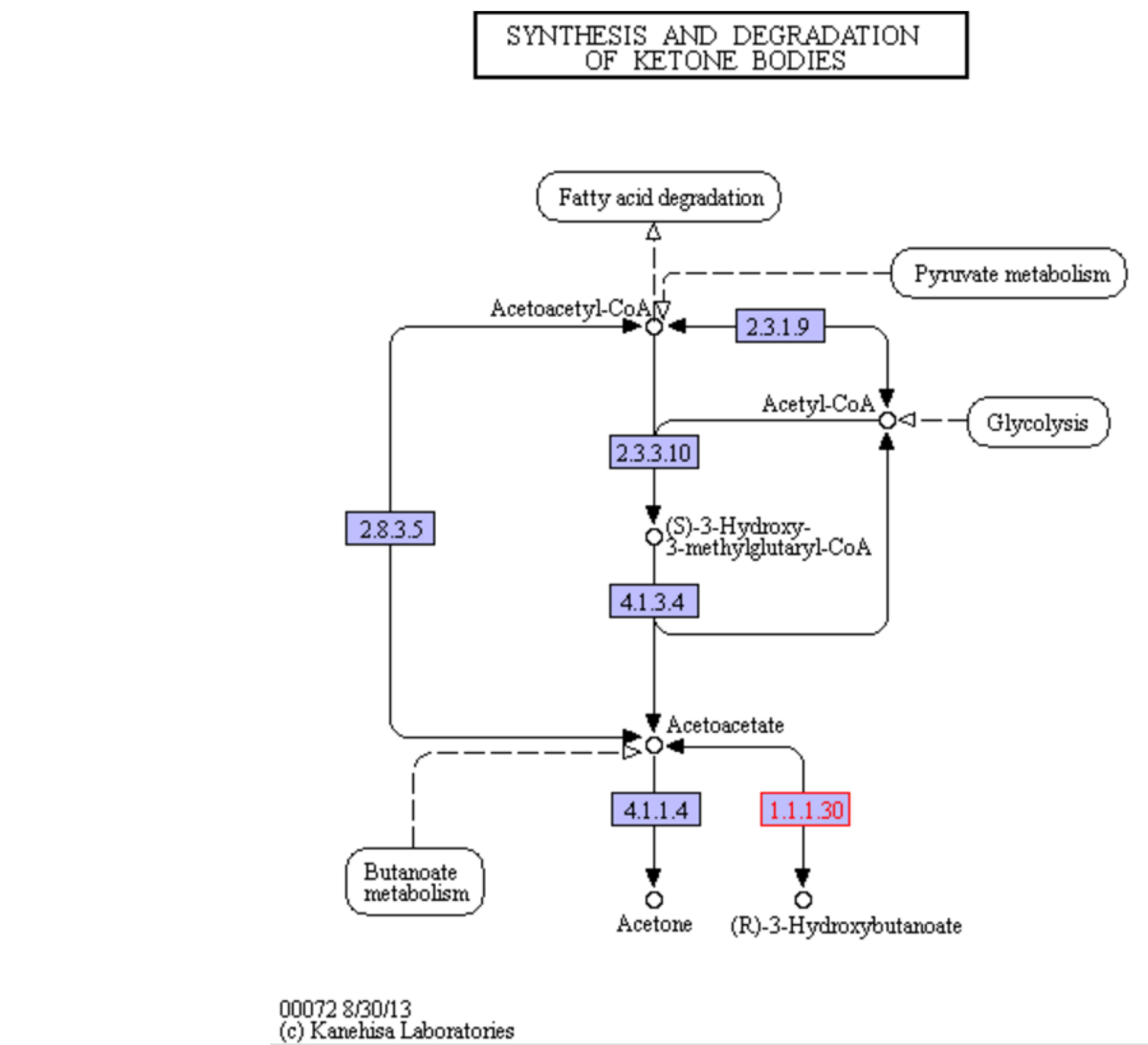


Figure 3: The enzymatic pathway of Ksed_RS05035, Hydroxybutyrate

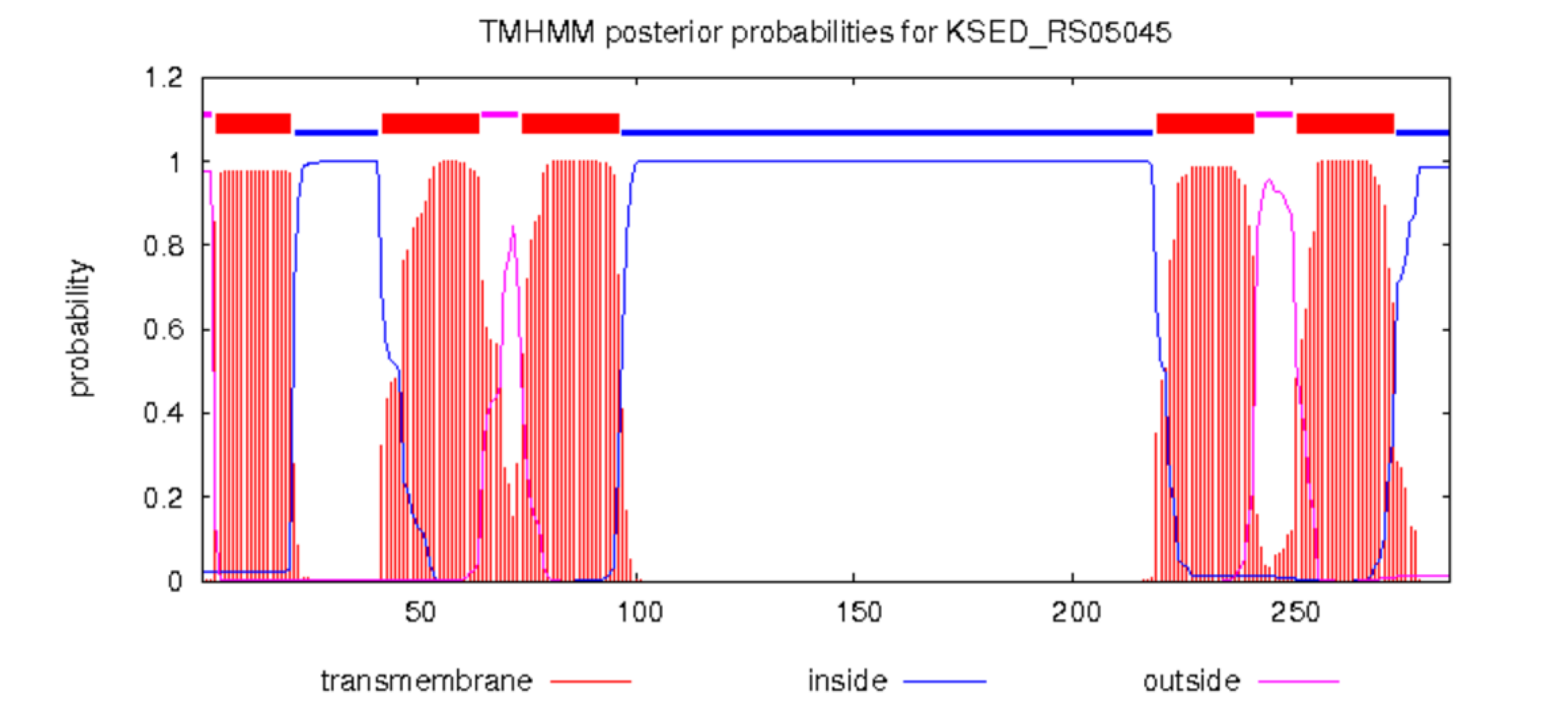


Figure 4 – A TMHMM result predicting 5 transmembrane helices for Ksed_RS05045

type II secretion system protein F [Kribbia dieselivorans]									
Sequence ID: WP_068401563.1 Length: 288 Number of Matches: 1									
Range 1: 3 to 288 GenPept Graphics									
Score	Expect	Method	Identities		Positives		Gaps		
329 bits(843)	4e-110	Compositional matrix adjust.	174/287(61%)		214/287(74%)		3/287(1%)		
Query 2		GHLLGATLGLLGLVWDSFW--PRPRARVTRTTQDRVSDLAVARVSGVTPGGLVLT				59			
		G+ LGA LG GL LW + W P PR +R D + D+AA A +G TP G++ +							
Sbjct 3		GLFLGALLGTGLFLVMACIAPQVEPRVG--SRGRMLDLRDEIAQAAGTGTATGTVLTS				61			
Query 60		AALGVTFVLTAGLTVLPAFAAGLITASGLPMAVLSGTSRRRRRLRDAMPDCVHVAS				119			
Sbjct 62		L V F+ V+P+AA FGLI+AGLP ++ G +R RR LRD WPD VD++S							
		VLVFAVFLFGVATVYKVPVLAACFLIAGLGP IIVIGRRHRRRLDDVDVONISS				121			
Query 120		GVRAGMALPEALQAVRGPEELRPAPAFAGFAHDYRATGRFHVSLDALKERLSDPVADRLV				179			
Sbjct 122		VRAG+ALPEALQ LVRGP ELRPAF F DYRATGRFH LD LKERLSDPVADRL+							
		AVRAGLALPEALQGVGVRLPFAEFGDYRATGRFHVHCDLKERLSDPVADRLV				181			
Query 180		ESLIRAREVGGSGLGVRLTSSFLREEAHRSEMEARQSVTVNAARLGLAAPVILVAML				239			
Sbjct 182		ESLR+AREVG+DLG +LRTLSSFLE+R+ TR+E+E RQ+HTVNAARL +AAPVILVAM+							
		ESLRVAREVVGTDGLSLRLTSSFLREDARLTAETRGVATVNAARLVAAPVILVAM				241			
Query 240		ATRGDSLAAYNSTTGALLVLGLGVSGVAYLMMRLGRTPQAEVRLR				286			
Sbjct 242		TRPDS+AAVY+TG LVL +GG ++V AY LHM+RGR+ +ERVLR							
		TTTRPDSMAAYGTGVLVLAIGGALTVAYQMLRLIGDLSRDEPR				288			

Figure 5 – Example of a BLAST hit for Ksed_RS05045

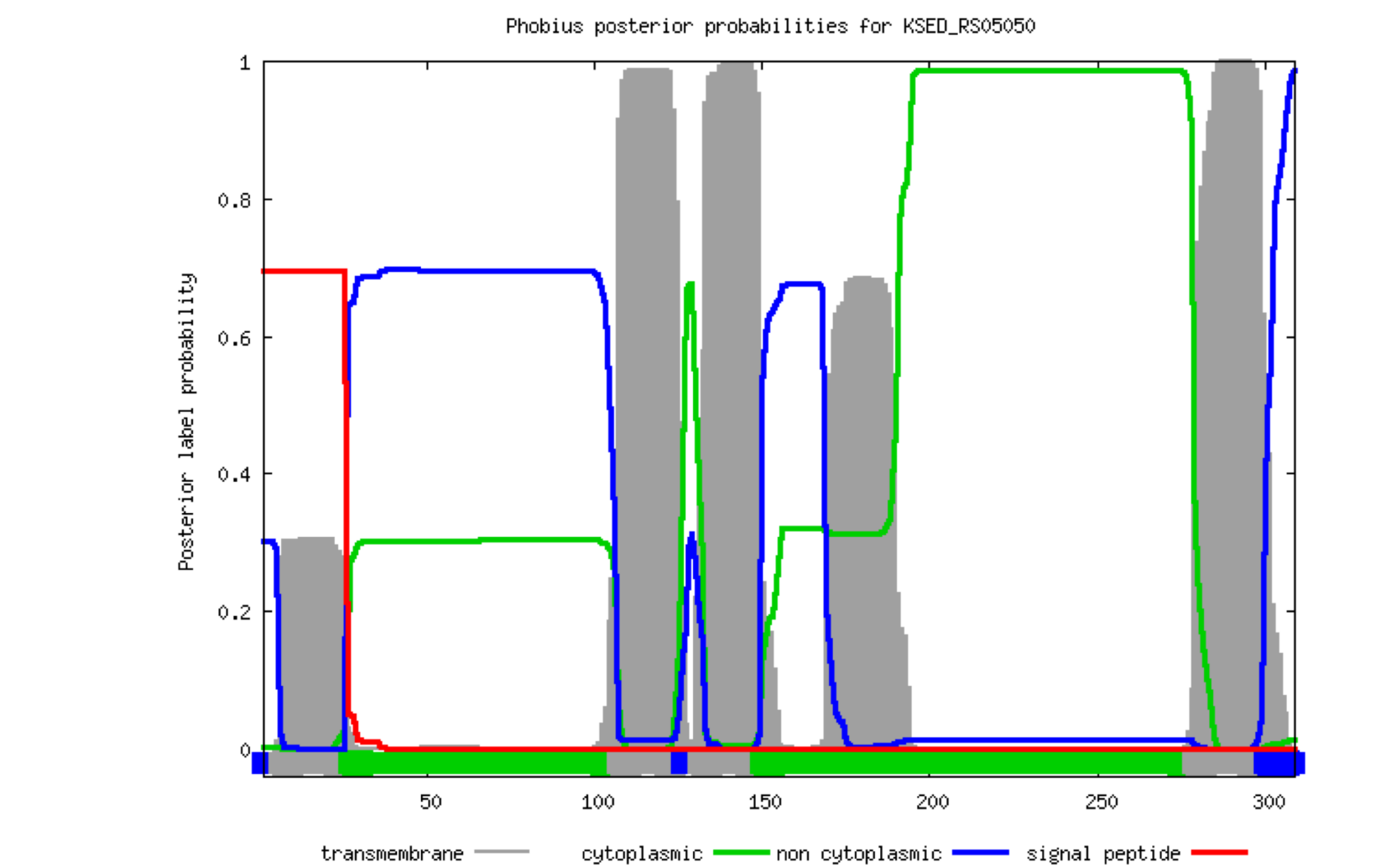
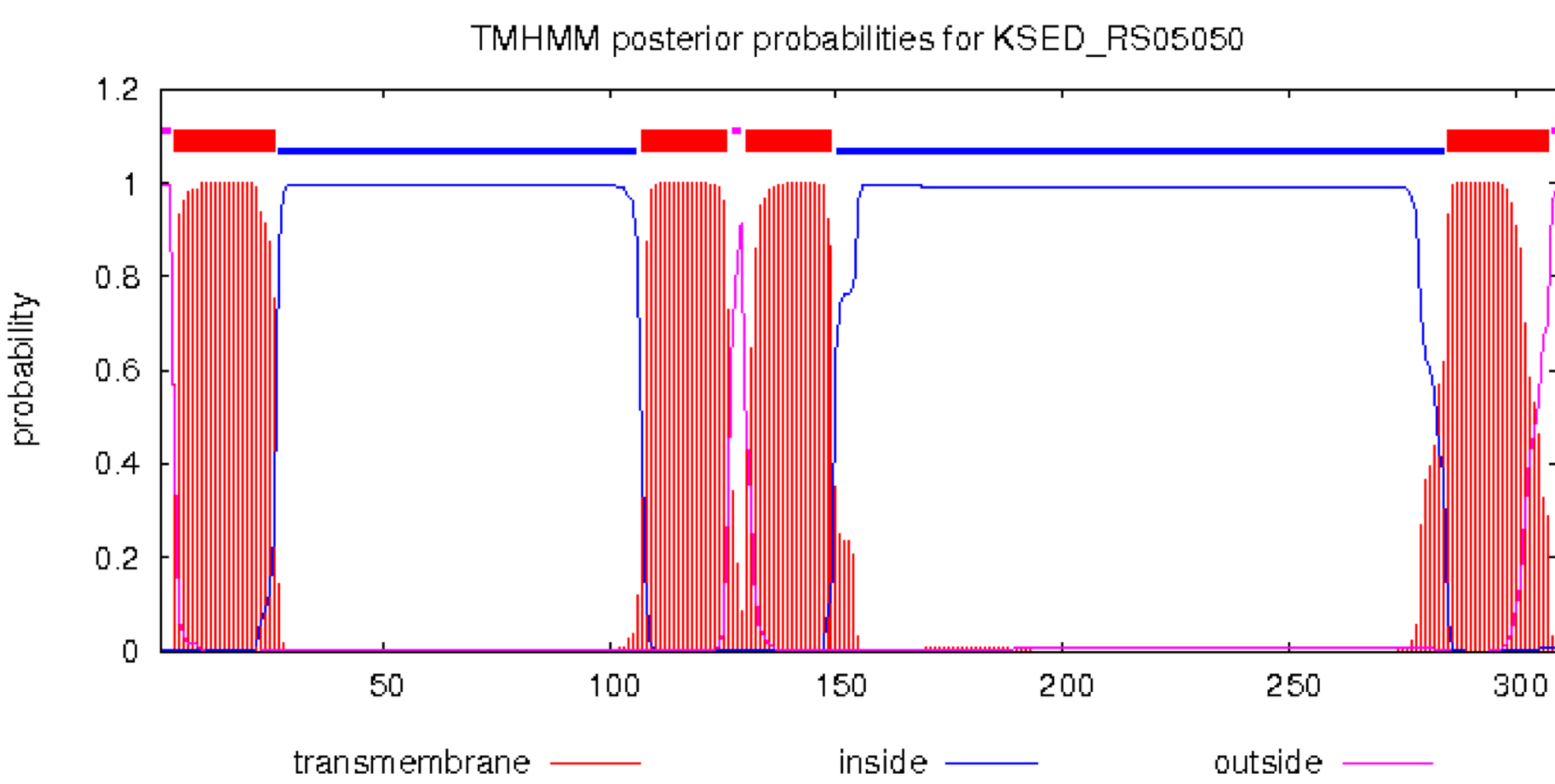


Figure 6 – The TMHMM result predicts that Ksed_Rs05050 has 4 transmembrane helices. Phobius confirms that there are transmembrane helices.

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 Product	Proposed Annotation
Ksed_RS05035	3-hydroxybutyrate dehydrogenase	3-hydroxybutyrate dehydrogenase
Ksed_RS05045	type II secretion system protein F	type II secretion system protein F
Ksed_RS05050	Flp pilus assembly protein TadB	Flp pilus assembly protein TadB

References

Folayan et al. (2018). *Kytococcus Sedentarius* and *Micrococcus luteus*: highly prevalent in indoor air and potentially deadly to the immunocompromised – should standards be set? Tropical Biomedicine, 35 (1) 149-160.
Sims et al. (2009). Complete genome sequence of *Kytococcus sedentarius* type strain (541T). Standards Genomic Sciences.12 -.20.

Acknowledgments

Supported by an NIH Science Education Partnership (SEPA) Award - R25OD010536
Special thanks to Dr. Stephen Koury, Dr. Rama Dey-Rao, and Brittany Mitchell