SEPA SCIENCE EDUCATION PARTNERSHIP AWARD
Supported by the National Institutes of Health

Annotation of the *Myxococcus xanthus* Genome from Locus Tags MXAN_RS00865 to MXAN_RS00900

Jacie Beebe, Natalie Drumm, Jadyn Gardner, Emma Kephart, Sean McKenna, Krista Smith, Joel VanSkiver, and Julie McLaughlin
Campbell-Savona Jr/Sr High School, Campbell, NY and the Western New York Genetics in Research and Health Care Partnership

Abstract

A group of 7 semi-consecutive genes from the microorganism *Myxococcus Xanthus* (MXAN_RS00865 – MXAN_RS00900) were annotated using the collaborative genome annotation website GENI-ACT. The GenBank proposed gene product name for each gene was assessed in terms of the general genomic information, amino acid sequence-based similarity data, structure-based evidence from the amino acid sequence, cellular localization data, and potential alternative open reading frames. The GenBank proposed gene product name 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 in the database.

Introduction

Myxococcus xanthus is a motile gram negative bacteria that forms elaborate biofilms. They belong to the delta class of the proteobacteria, and have a relatively large genome. They are known for their social behaviors including swarming, predation, and development (Wang *et. al*. 2011).

This soil bacterium's large genome can be partially attributed to horizontal gene transfer and certainly allows them to exhibit their predatory behavior (Wang *et. al.* 2011). *Myxobacteria* tend to have multicellular structures, like fruiting bodies when deprived of nutrients, and are characterized by cellulosic and proteolytic activities (Keane and Berleman 1). Because of this, they perform important roles in the global carbon cycle and influence the flow of carbon in microbial communities. Their motility, both social (S) and adventurous (A), allows *Myxococcus xanthus* to perform the multicellular activities of scouting, branching, rippling, and aggregating into fruiting bodies preceding sporulation (Keane and Berleman 2016). According to Wang *et. al.*, *Myxococcus xanthus* cells use degradative enzymes, antibiotics, and bioactive compounds to lyse target prey (2016). They are considered predators because they can then use the nutrients released by this lysis to grow and support metabolic activities. So, *Myxococcus xanthus* are intriguing subjects of study both genetically and behaviorally.

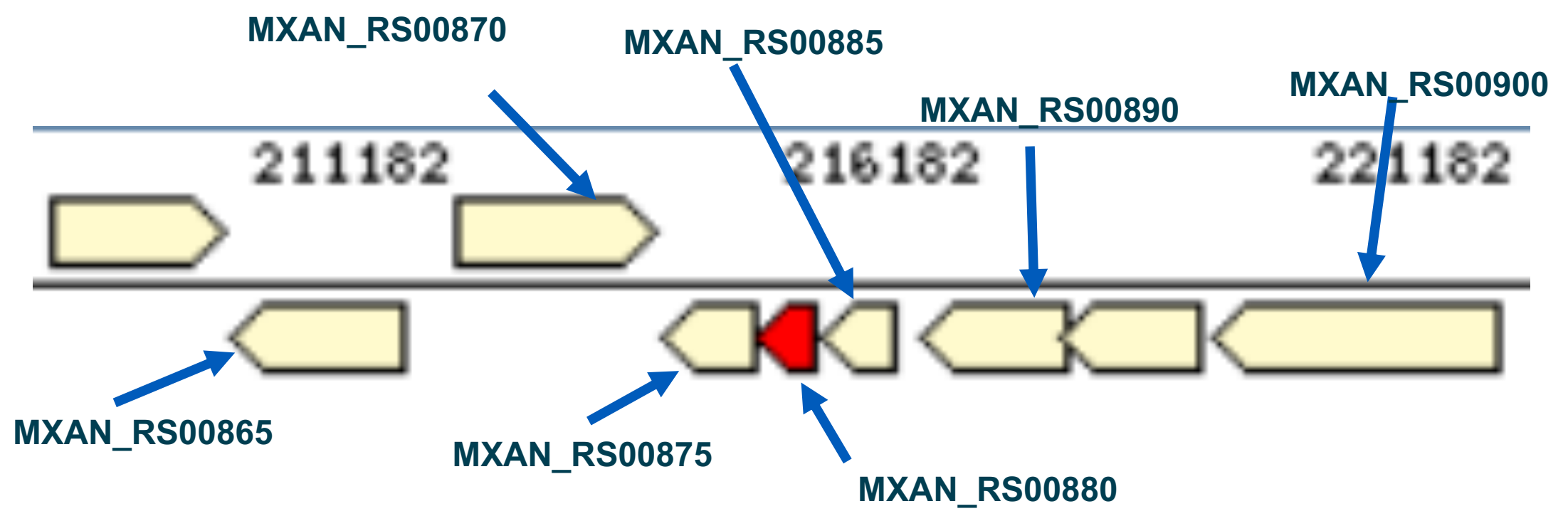

Figure 1. The locus tags and relative position of the genes under investigation in this research

Methods

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

Modules	Activities	Questions Investigated
Basic Information	DNA Coordinates and Sequence, Protein Sequence	What is the sequence of the gene and protein? Where is it located in the genome?
Sequence-Based Similarity	Blast, CDD, T-Coffee, WebLogo	How similar is the protein under investigation to other proteins in GenBank?
Structure-Based Similarity	TIGRfam, Pfam, PDB	What functional domains are present in the protein under investigation?
Cellular Localization	Gram Stain, TMHMM, SignalP, LipoP, Psortb, Phobius	Is the protein under investigation located in the cytoplasm, secreted, in the periplasm or embedded in the cell membrane or cell wall?

Results

MXAN_RS00865

The computer pipeline proposed product of this gene was a tetrahydrofolate ligase. 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 cytoplasmic localization of the amino acid sequence.

MXAN_RS00870:

The computer pipeline proposed product of this gene was a PAS domain S-box protein. This 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 intercellular localization of the amino acid sequence within the cytoplasmic membrane based on the PSORT-B results.

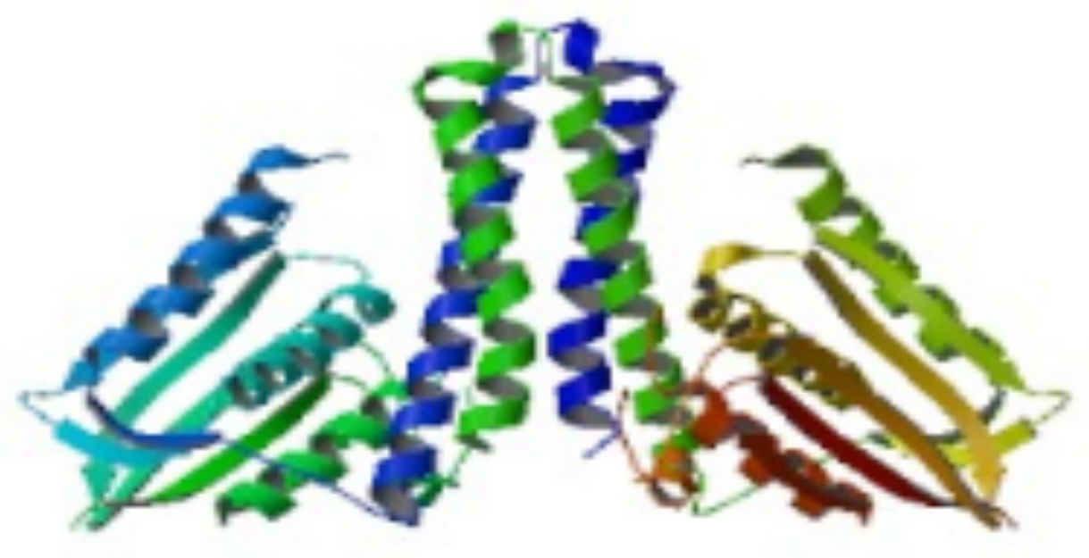

Figure 2. Inactive structure of histidine kinase with a similar structure to that of the PAS domain S-box protein.

MXAN_RS00875:

The computer pipeline proposed product of each of these genes was a sulfite exporter TauE/SafE family protein. This was supported by the top BLAST hits for the amino acid sequences, the presence of well-curated functional domains within the amino acid sequences, the cellular location of the amino acid sequences, and the enzymatic function of the amino acid sequences.

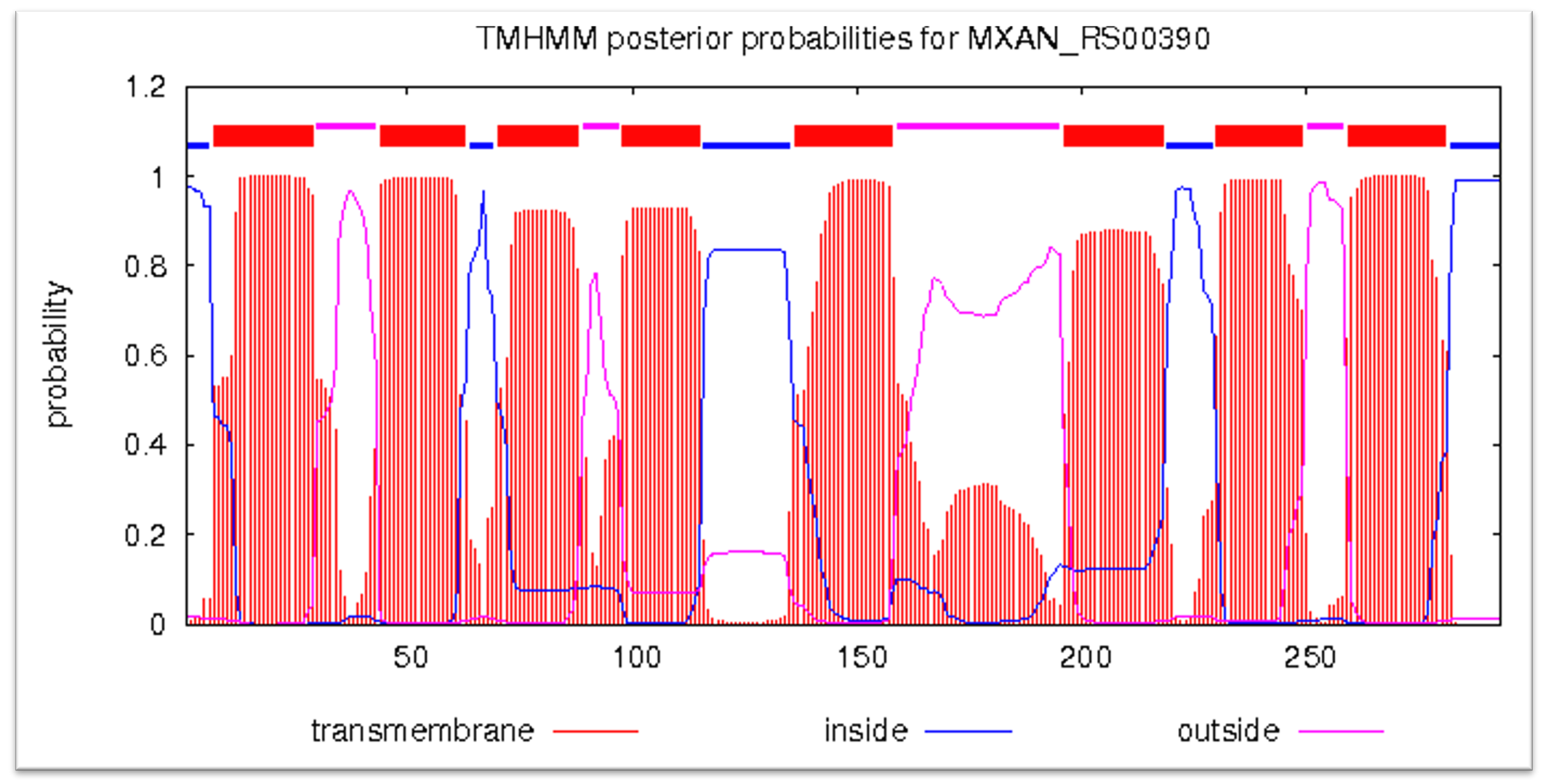

Figure 3. TMHMM Results for MXAN_00875 showing 8 transmembrane helices.

MXAN_RS00880:

The initial proposed product of this gene was rhodanese-like domain-containing protein found in the organism *Myxococcus virescens*. This 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. From CDD the top hit for COG number was COG0607 with the name PspE Pfam results further supported that the gene was in a Rhodanses-like domain with a clan number of CL0031. Due to the 34% identity and very low e-value of 3.63837E-5 it is likely that the query protein has a structure similar to this protein Entity 1 containing Chain A,B.

MXAN_RS00885:

The initial proposed product of this gene was a membrane component Zn-dependent hydrolase. This 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.

Figure 4. Crystal structure of ETHE1 from Myxococcus xanthus.

MXAN_RS00890:

The initial proposed product of this gene was an acetoacetate metabolism regulatory protein AtoC. 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. The number 1 blast hit was an acetoacetate metabolism regulatory protein AtoC and it obtained an E-value of 0.0. From CDD the top hit for GOG number was COG3829 with the name RocR. Pfam results further supported that the gene was in a Sigma-54 interaction domain with a clan number of CL0023.

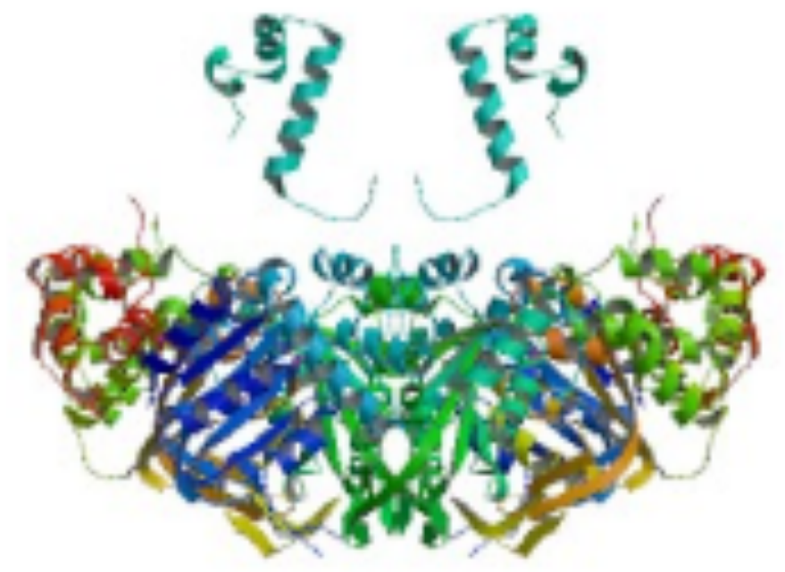

Figure 5. Crystal structure of a sigma54-activator suggests the mechanism for the conformational switch necessary for sigma54 binding.

MXAN_RS00900:

The computer pipeline proposed product of this gene was a AsmA family protein. This 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 transmembrane location of the amino acid sequence.

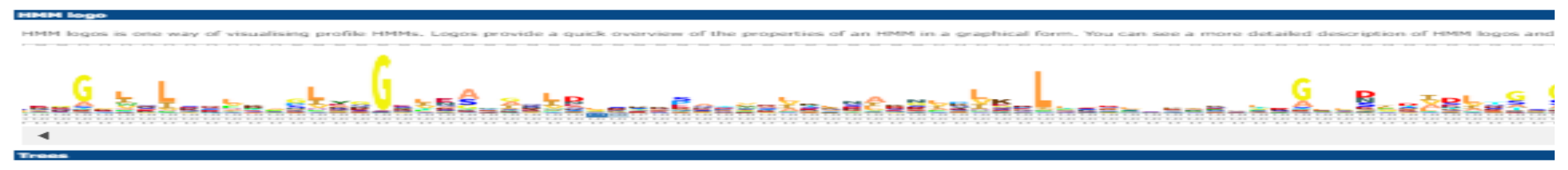

Figure 6. HMM logo of PF05170 matching MXAN_RS00900 showing conserved amino acids.

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.

Locus Tag	Pipeline Annotation Product Name	Proposed Annotation	Changes Proposed?
MXAN_RS00865	tetrahydrofolate ligase	tetrahydrofolate ligase	No
MXAN_RS00870	PAS domain S-box protein	PAS domain S-box protein	No
MXAN_RS00875	sulfite exporter TauE/SafE	sulfite exporter TauE/SafE	No
MXAN_RS00880	rhodanese-like domain-containing protein	rhodanese-like domain-containing protein	No
MXAN_RS00885	membrane component Zn-dependent hydrolase	membrane component Zn-dependent hydrolase	No
MXAN_RS00890	acetoacetate metabolism regulatory protein AtoC	acetoacetate metabolism regulatory protein AtoC	No
MXAN_RS00900	AsmA Family Product	AsmA Family Product	No

References

Wang, J., Hu, W., Lux, R., He, X., Li, Y., & Shi, W. (2011). Natural transformation of Myxococcus xanthus. *Journal of bacteriology*, *193*(9), 2122–2132. doi:10.1128/JB.00041-1

Keane R, Berleman J. (2106). The predatory life cycle of Myxococcus xanthus. *Microbiology,* 162(1):1-11 doi:10.1099/mic.0.000208

Acknowledgments

Supported by an NIH Science Education Partnership (SEPA) Award - R25GM129209