Annotation of the *Moraxella catarrhalis* Genome at Locus Tags MCR_RS00145, MCR_RS00800, MCR_RS805, and MCR_RS810

Danya Flood, Yohannes Kesete, Sadnan Nawz, Tasfia Rashid, Fernando Vega, Kira Mioducki & Dawn Weihrich Research Laboratory High School, Buffalo Public Schools (BPS#366) and the Western New York Genetics in Research and Health Care Partnership

Abstract

Three genes located within the *Moraxella catarrhalis* genome 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 and structure-based evidence from the amino acid sequence. The GenBank proposed gene product name for each revealed a protein function related to transport. Results 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 database.

Introduction

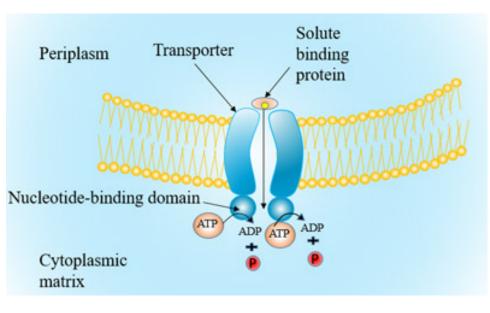
Moraxella catarrhalis is a non-motile bacterium that spreads rapidly on its host. Colonies of *M. catarrhalis* may have a rough surface, are pinkish-brown in color. This bacterium causes otitis media, a middle ear. It can also cause bronchitis, which inflames the lining of bronchial tubes that carry air to and from the lungs, making breathing difficult (Racaniello, 2019). This bacteria adheres and colonizes the respiratory tract by infecting mucous membranes and epithelial cells.

The three locus tags annotated are all predicted to function as transport proteins. Specifically, MCR_RS00145 is an iron transporter, associated with regulating iron levels at the cellular level. These proteins are located in the cell membrane and aid the cell in transporting various substances into and out of the cell, such as amino acids and ions like methionine and iron. They require ATP to transport substances across the membranes, as ATP-binding cassette (ABC) transporters utilize the energy from ATP and consequent hydrolysis to transport various substances across cellular membranes. These ABC transporters contain nucleotide-binding domains and bind DNA/RNA. The two significant domains in this process are: ATP binding and transmembrane domains.

Figure 1. Diplococcus shape of Moraxella catarrhalis bacteria.



Figure 2. Proteins belonging to the ABC (ATP-binding cassette) superfamily function as transporters that regulate levels within the cell.



Methods

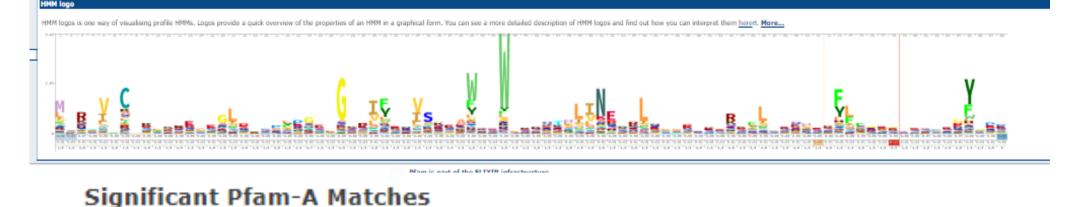
Modules of the GENI-ACT (http://www.geni-act.org/) were used to complete *Moraxella catarrhalis* genome annotation. The modules are described below:

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?
Final Annotation	Evaluate data from all modules	Has the gene been correctly called by the pipeline annotation?

Results

MCR RS00145:

The computer pipeline proposed product of this gene was an iron transporter. BLAST hits revealed organisms with similar pipeline proposed products, including Psychrobacter arcticus with a gene product name: probable Fe trafficking protein. This BLAST hit score was 142 bits with an e-value of 1e-44. Significant COG hits suggested alignment to a gene for cluster biosynthesis and repair proteins, along with transport for metabolism. When WebLogo was examined, a few sections of high conservation were evident throughout its length, with a lack of conservation at the N-terminal. PDB results show a significant hit of 1T07: Entity 1 containing Chain A Crystal Structure of Conserved Protein of Unknown Function PA5148 from Pseudomonas aeruginosa. The Pfam results show this protein as being in the family of iron trafficking and the HMM logo shows a highly conserved tryptophan residue.



Fa	mily	Description	Entry type	Clan	Sta
Iron	traffic	Bacterial Fe(2+) trafficking	Family	n/a	7
HMM MATCH IPP ISEQ	++vfc+k++ 79*******	keaegLdfpplpgelGkrifenvskeawqeWlkkqtmLiNek ++++ L++pp+p+++G++i +++s +aw++Wl+ qtmLiNek ************************************	<pre>k+L+++dp+a+k+L+eq ek k++++++dp+a+k+L+eq ek k+++++++++++++++++++++++++++++++++++</pre>	(fl ++++ekp+	gy+p **98

Figure 3. Pfam alignment for MCR RS00145 and HMM with significant tryptophan residue.

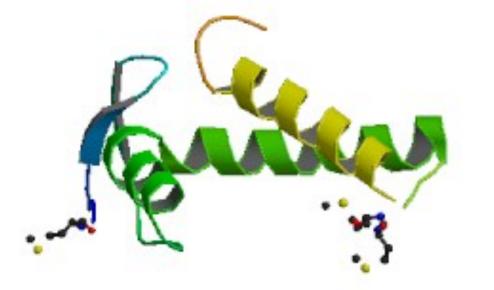


Figure 4. MCR RS00145 PDB result showing a 1T07: Entity 1 containing Chain A Crystal Structure of Conserved Protein of Unknown Function PA5148 from Pseudomonas aeruginosa.

Acknowledgments



Thank you to Dr. Stephen Koury, Dr. Rama Dey-Rao, Dr. Sandra Small & Jonathan E. Bard for your invaluable support and expertise.

MCR_RS00800: The computer pipeline proposed product was methionine ABC transporter substrate binding protein MetQ. The gene sequence annotated is 276 amino acids in length and located between coordinates 188301 and 189131. BLAST hits suggest alignment with gene products relating to Dmethionine binding lipoprotein and MetQ. Significant COG hits indicated alignment to NIpA, an ABC type metal ion transport system, periplasmic component/surface antigen. PDB results reveal a significant hit of 4YAH: Entity containing Chain X, Crystal Structure of the Methionine Binding Protein, MetQ.

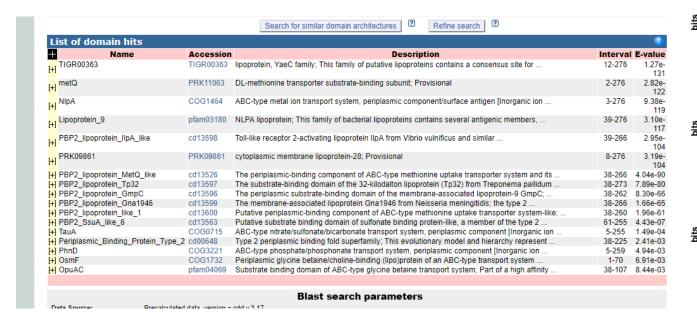


Figure 6 (above). Results for Conserved Domain Database for MCR_RS00800 show domains aligned to predicted function including ABC-type Methionine transporter.

Figure 7 (right). WebLogo results for MCR_RS00800 show well conserved residues throughout, with gaps at the N-terminal.



MCR RS00800 Protein Data Figure 5. Bank top result displaying the crystal structure of the Methionine Binding Protein, MetQ, 4YAH.

MCR_RS00810:

The proposed initial product was D-Methionine ABC transporter, ATP-binding protein. The gene annotation is between 190002 and 191048 coordinates on the genome. It is 1047 nucleotides in length and 848 amino acids. The scores of the top BLAST were 456 bits and 453 bits respectively. The top two significant BLAST hits in SwissPro revealed the same protein function: Methionine import ATPbinding protein MetN. Significant COG alignment also showed ABC-type methionine transport system, ATPase component. T-coffee alignment and WebLogo showed well conserved residues across various species. PDB results indicate 3DHW: entity 2 containing Chain C, D, G, H Crystal structure of methionine importer MetNI.

 Image: Second $= \frac{1}{2} - \frac$

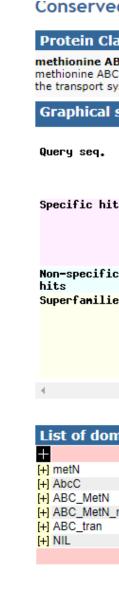


Figure 8. Conserved Domain Database results for MCR RS00810 reveal consistent alignment to ATP-binding cassette domains, specifically D-methionine transport.

MetNI.

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 Ta

MCR_RS0 MCR_RS

MCR_RS00

IMAGE SOURCE: A transport machine -- ATP-binding cassette. (n.d.). Retrieved May 1, 2019, from https://www.cusabio.com/c-20790.html.

(n.d.). Retrieved from https://www.microbiologyinpictures.com/bacteria photos/moraxella catarrhalis photos/moraxella catarrhalis 01.html

buffalo.edu



ed domains on [gi 295920099 gb ADG60450]						
ass	ification				?	
	nsporter ATP-bindi		in (domain architecture ID 11485231) tN is part of the ABC transporter complex MetNIQ that is involved in methionine import; responsible fo	r energy co	upling to	
sun	nmary 🔲 Zoo	om to residu	e level show extra options »		2	
ts	ATP binding site Halker A/P-loop	50 ABC Q-loop/	100 150 200 250 300 transporter signature motif D-loop Halker B H-loop/switch region metN	348		
	×		AbcC			
С			ABC_HetN_methionine_transporter NIL ABC_tran ABC_MetN	•		
es		-	ITPase superfamily NIL superfamily C_ATPase superfamily			
			ABC_MetN superfamily			
			AbcC superfamily			
			metN superfamily			
			Search for similar domain architectures ? Refine search ?		÷.	
nai	n hits				2	
Na	ame	Accession PRK11153 COG1135	Description DL-methionine transporter ATP-binding subunit; Provisional ABC-type methionine transport system, ATPase component [Amino acid transport and metabolism];	Interval 1-347 1-346	E-value 0e+00 2.48e-180	
_meth	nionine_transporter	cd03258 pfam00005	D-methionine ABC transporter, ATP-binding protein; Members of this family are the ATP-binding ATP-binding cassette domain of methionine transporter; MetN (also known as YusC) is an ABC transporter; ABC transporters for a large family of proteins responsible for translocation This domain is found at the C-terminus of ABC transporter proteins involved in D-methionine	1-347 1-237 29-173 271-344	6.48e-172 1.47e-147 3.77e-47 9.52e-16	

Figure 9. Protein Data Bank results displaying crystal structure of methionine importer



g	Pipeline Annotation Product Name	Proposed Annotation	Changes Proposed?
00145	Iron transporter	Iron transporter	No
00800	Methionine ABC transporter substrate-binding protein MetQ	Methionine ABC transporter substrate-binding protein MetQ	No
0810	D-methionine ABC transporter, ATP-binding protein	D-methionine ABC transporter, ATP-binding protein	NO

References

First Med Urgent Care OKC & Edmond. (2015, Sept). Urgent Care Reports: Signs of Ear Infections for Children and Adults. Retrieved from http://firstmedok.com/2015/02/25/signs-of-ear-infections-for-children-and-adults/

Racaniello, V. (2009, May). Viruses and the Respiratory Tract. [Blogpost]. Retrieved from http://www.virology.ws/2009/05/21/viruses-and-the-respiratory-tract/

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

Jacobs School of Medicine and Biomedical Sciences