

Annotation of the *Moraxella catarrhalis* Genome from Locus Tags MCR_RS00140 to MCR_RS00145

Danya Flood, Bre'Jon Sanders, Mohaddeseh Golzar, Mary Manhong, Kira Mioducki, and Dawn Weirich
 Research Lab Program for Bioinformatics and Life Sciences, Buffalo, NY and the Western New York Genetics in Research And Health Care Partnership



Abstract

Two consecutive genes from the microorganism *Moraxella catarrhalis* (MCR_RS00140 and MCR_RS00145) 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 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 using the genetic databases.

Introduction

Moraxella catarrhalis is an emerging human respiratory tract pathogen. This microorganism is a common cause of childhood otitis media, middle ear infection, and exacerbation of chronic obstructive pulmonary disease in adults. It is found in mucous membranes of the respiratory tract and epithelial cells (MD-Health, 2012). *Moraxella catarrhalis* is a non-motile bacteria that spreads rapidly on its host. It has a rod shape form and is a diplococcus, a round bacterium that typically occurs in the form of two joined cells (Brophy, 2017). *Moraxella catarrhalis* transmits through breathing and coughing from person to person in a hospital setting, and they occur more in people with complications of the respiratory system. *Moraxella catarrhalis* is resistant to penicillin, amoxicillin and ampicillin (Soklich, 2015).

The *Moraxella catarrhalis* gene with the locus tag MCR_RS00140 that was annotated is involved in CRISPR. This is a type of adaptive immune system found in this bacteria. Type I-F CRISPR adaptive immune system in *Moraxella catarrhalis* consists of two CRISPR loci and six CRISPR-associated (Cas) genes (Hille et al., 2016). Type I-F systems rely on a CRISPR RNA (crRNA)-guided surveillance complex (Csy-complex) to bind foreign DNA and recruit a nuclease (i.e., Cas2/3) for target degradation (Cooper et al., 2018).

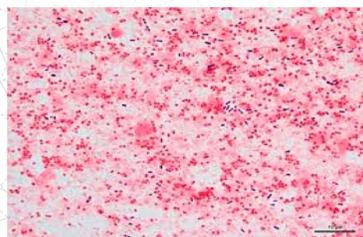


Figure 1-*Moraxella catarrhalis*
<http://microbe-canvas.com/Bacteria.php?p=1195>

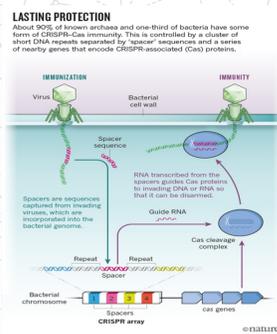


Figure 2 Explanation of how CRISPR-Cas immunity works. (Ledford, 2017)

Methods

Modules of the GENI-ACT program (<http://www.geni-act.org/>) were used to complete *Moraxella catarrhalis* 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 5- Structure-Based Evidence	Pfam, PDB	Are there functional domains in my protein?

Results

MCR_RS00140:

The initial proposed product of this gene by GENI-ACT was type I-F CRISPR-associated endoribonuclease Cas6/Csy4. The BLAST database indicated that this gene product's top hit was a multi species – type I-F CRISPR-associated endoribonuclease Cas6/Csy4 also found in the organism *Acinetobacter*. The BLAST indicated 66% positive alignment (see figure 4). The PFAM database indicated a protein family name of Cas_Csy4 and a clan name was RAMPS-CAS 5-like. Multiple sequence alignment, as shown in WEBLOGO, had good conservation of amino acids between residues 1 -141. However, from 142 and 231 there are significant gaps especially toward the C-terminal (see figure 6).

MCR_RS00145:

The initial proposed product of this gene by GENI-ACT was an iron transporter. The initial BLAST test showed the top hit as being an oxidative damage protective protein. But, by doing the PFAM and COG tests, the results support the initial proposed gene name. The PFAM test indicated an Fe(2+) trafficking protein family, and the COG revealed results for Fe-S cluster biosynthesis and repair protein YggX [inorganic ion transport and metabolism, posttranslational modification, protein turnover, chaperones]. Most of the letters displayed in Weblogo are three bits tall. There are some gaps between residues 3-9 and 97-99 (see figure 7). Most of the tall letters are black. This may indicate that the amino acids are hydrophobic and, therefore, supports the hypothesis that this protein may be located in a membrane and aid in trafficking of iron.



Figure 3 The locus tags and relative position of the genes under investigation in this research.

MULTISPECIES: type I-F CRISPR-associated endoribonuclease Cas6/Csy4 [*Acinetobacter*]
 Sequence ID: WP_009584811.1 Length: 203 Number of Matches: 1
[See 1 more title\(s\)](#)

Score	Expect	Method	Identities	Positives	Gaps
216 bits(551)	2e-68	Compositional matrix adjust.	109/203(54%)	135/203(66%)	7/203(3%)

Range 1: 1 to 203 GenPept Graphics Next Match Previous Match

Query 1 MKFYQETLIPQADISPHAVNSKLYTQLHAFVEKDEQDKTIVGVSPQVRTIADKKA 60
 MK+Y EITLI D S VNSKLYTQLHAF E Q D + K I+G+SFPQVR K I
 Sbjct 1 MKYYEITLIESIDPSMFEVNSKLYTQLHAFAEQIDTRGKIIFGISFPQVRIHQKNI 60

Query 61 YLGGKLRVFAPTQALSALNLDKRLHLDVTHISSRSVDP-DIKGVANVYRAIPKMTL 119
 +LG K+RVFAP+E L LNL K LE L DV+H+ R VP I GVA+YAR PHTL
 Sbjct 61 FLGSKRVFAPSELEQLNLSKLERLDVYVHTQPREVPKTKITGVYHFRINPKMTL 120

Query 120 DERITHQAKRHGVPYHKAERFEMKYEQSLVYPHIQLTSKSHRS-----VPLHIGKQA 173
 +ERI HQA+R + + + + F+ Y EQ + P+ L S S R Y L+IGK
 Sbjct 121 EERIAHQARRSISFDEAQHFQKQYSEQPEIVPVLSKLSLTKREIQIRYRLVIGKSV 180

Query 174 TEVLTDFRGTGVLRSRTSVPFEF 196
 E DG+FTYGLSRT++VPEF
 Sbjct 181 AEEANDGKFGTYGLSRTATVPEF 203

Figure 4 – Top NR BLAST hit for Locus Tag MCR_RS00140

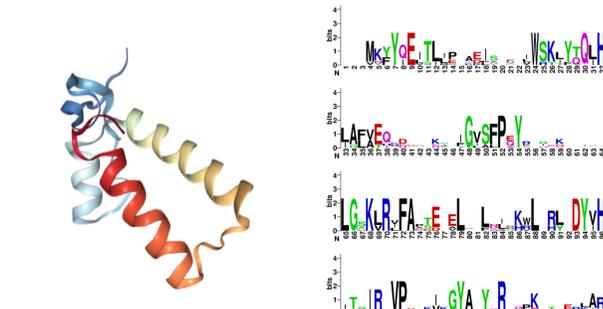


Figure 5 (above) - Crystal structure of Locus Tag MCR_RS00145 obtained from Protein Data Bank

Figure 6 (right) - Weblogo Results from Locus Tag MRC_RS00140 which shows incomplete alignment at the C-terminal.

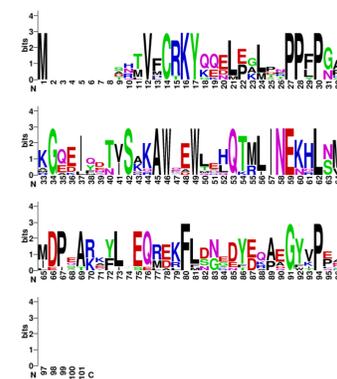


Figure 7 – Weblogo Results from Locus Tag MRC_RS00145 which indicates highly conserved amino acid sequence except for the C-terminal.

Conclusion

The Geni-Act proposed gene products were mostly consistent with the Geni-Act proposed gene product. However, there was some discrepancy with Locus Tag MCR_RS00145. There was evidence of the protein being an iron transporter and an oxidative damage protective protein. A possible explanation for these results may be that iron transporters assist in redox reactions and prevent oxidative stress. Further biochemical testing to confirm our suspected annotation may be needed.

The proposed function of the locus tag MCR_RS00140 is an endoribonuclease, which is a CRISPR-associated enzyme that degrades foreign DNA and protects the bacteria from viruses and other foreign genomes (Carte, J. 2008).

Gene Locus tag	Geni-Act proposed gene product	Proposed Annotation
MCR_RS00140	I-F CRISPR-associated endoribonuclease Cas6/Csy4	Type I-F CRISPR-associated endoribonuclease Cas6/Csy4
MCR_RS00145	Iron Transporter	Iron Trafficker

References

Brophy Martinez, K. (2017, July). MLAB 2434: Microbiology. [Powerpoint Slides]. Retrieved from <http://slideplayer.com/slide/5794094/>

Carte, J., Wang, R., Li, H., Terns, R. M., & Terns, M. P. (2008). Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes & Development*, 22(24), 3489–3496. <http://doi.org/10.1101/gad.1742908>

Cooper LA, Stringer AM, Wade JT. 2018. Determining the specificity of Cascade binding, interference, and primed adaptation in vivo in the *Escherichia coli* type I-E CRISPR-Cas system. *mBio* 9:e02100-17. <https://doi.org/10.1128/mBio.02100-17>

Hille, F. and Charpentier, E. (2016). CRISPR-Cas: biology, mechanisms and relevance. *Phil. Trans R. Soc. B*. 371:20150496. Retrieved from <http://dx.doi.org/10.1098/rstb.2015.0496>

Ledford, H. (2017, Jan). Five big mysteries about CRISPR's origins. Retrieved from <https://www.nature.com/news/five-big-mysteries-about-crispr-s-origins-1.21294#auth-1>

MD-Health. (2012, Dec.) Bronchitis, Is it Contagious? Retrieved from <http://www.md-health.com/Is-Bronchitis-Contagious.html>

Soklich, T. (2015, Oct). Does being cold make you sick? Retrieved from <https://sites.psu.edu/siowfa15/2015/10/05/does-being-cold-get-you-sick/>

Vidyasagar A. (2018, April). What is CRISPR? Retrieved from Live Science <https://www.livescience.com/58790-crispr-explained.html>

Acknowledgments

Special thanks to Dr. Rama Dey-Rao, Dr. Stephen Koury, and Dr. Sandra Small for their assistance with this project.

This work was supported by the National Science Foundation ITEST Strategies Award Number 1311902