

Annotation of the *Listeria Monocytogenes* Genome at Locus Tag LMOSLCC2372_1930 and LMOSLCC2372_2300

Vamsikrishna Ratnakaram, Christian Cowart, Student Name and Mrs. Sangeeta Gohkale
 Buffalo Engineering Awareness for Minorities at UB and the Western New York Genetics in Research Partnership



Abstract

The reason for our research was to investigate virulence in food-borne illness thus we picked proteins that were related to this cause. Our objective was to propose an annotation for a group of consecutive 2 genes from the microorganism *Listeria monocytogenes serotype 1 2c SLCC2372 FR733648* with the locus tags (LMOSLCC2372_1930) and (LMOSLCC2372_2300). We annotated them using the GENI-ACT public software and gene annotation program which categorized them based on things like enzymatic function and sequence-based similarity. Each Database and search driver corroborated and continued to validate our results, until eventually we could make a final annotation for our gene. LMOSLCC2372_1930 was finally annotated as Hemolysin III while LMOSLCC2372_2300 was annotated as Hemolysin.

Introduction

Listeria monocytogenes is a pathogenic gram-positive bacteria that causes approximately 260 deaths a year. It has a 20-30% fatality rate and is considered the third most dangerous foodborne illness causing more deaths than *Salmonella*. It is especially prevalent in pregnant women. If it spreads to the nervous system it can cause serious meningitis and encephalitis.

Hemolysin III was the first virulent factor discovered in *Listeria monocytogenes* in the 1980s. Hemolysin involves in the disruption of phagosomal membranes and allows the bacteria to escape destruction by enzymes in the white blood cells. This is key to the bacteria's survival. In my research I found that Hly-3 was a transmembrane protein that plays a key role in the breakdown of phospholipids thus the cellular membrane itself. (Vázquez-Boland et. al 2001)

Listeria monocytogenes uses hemolysin like CBS Domain proteins in their defensive mechanism by damaging host blood cells. It is crucial for their survival. Capacity of this bacterium to secrete such proteins is important for their virulence but not their entry in the host. (Mickaël Desvaux and Michel Hébraud, Sept 2008) (M. Kuhn, S. Kathariou, W Goebel, Jan 1988) The protein is located in transmembrane and composed of N-terminal enzymatic domain responsible for functions like metabolic enzymes, kinases, and channels

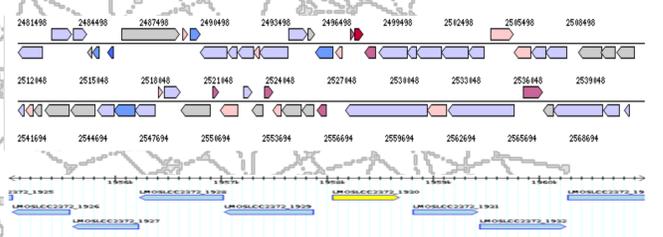


Figure 1. The locus tags and relative position of the genes under investigation in this research

Methods

Modules of the GENI-ACT (<http://www.geni-act.org/>) were used to complete *Kyococcus 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

LMOSLCC2372_1930:
 The initial proposed product of this gene by GENI-ACT was Hemolysin III. The two top nr Blast hits are Hemolysin III and Hemolysin [*Listeria monocytogenes*]. The top COG hit is: Predicted membrane channel-forming protein YqfA, hemolysin III family [Intracellular trafficking, secretion, and vesicular transport]. The top TIGRfam hit is: hlyIII: channel protein, hemolysin III family. The top Pfam hit is: Haemolysin-III. All of these findings support the fact that this protein is Hemolysin III.

RecName: Full=Hemolysin-3; AltName: Full=Hemolysin III; Short=Hly-III
 Sequence ID: P54176.1 Length: 219 Number of Matches: 1

Range 1: 13 to 214 GenPept Graphics

Score	Expect	Method	Identities	Positives	Gaps
232 bits(592)	9e-77	Compositional matrix adjust.	115/202(57%)	153/202(75%)	0/202(0%)

Query 9 KEEVANAITHGIGFISIPALVLLIIFAAGKDNPLYLTSFLYIGSIMLLYICSTLLHSF 68
 KEE+ANAITHGIG ILSIPAL+LII A+ + +F +YG+S+ LLY+ STLLHS
 Sbjct 13 KEEIANAITHGIGAILSIIPALILIIHASKHGTSASAVATVYGVSMFLYLSTLLHSI 72

Query 69 KPCKARTIFNIMDHAAYVLIAGSYTFVFLITIGTGLWLVGVIWLAIGLIYKIFMT 128
 K +F I+DH+AY+LIAG+YTF+LIT+G LGWL +IW LAI GI+KIF
 Sbjct 73 HHPKVEKLTILDRSAIYLLIAGTYFPLITLIRGLPWILLAIWLAIGGIIFKIFV 132

Query 129 GKLLKLLSTVSYLLMGMVFAIKPIYAGLIPFGWLLTAGGIMTFVGVAVFISIPRVPMH 188
 + ST Y++MGW++ AIRFLY LI GF LL GGI+++VGA+P+ ++P+H
 Sbjct 133 RRFIKASTLCYIIMGWLIIVAIKPIYENLITGHGFSILLAGSILVSGAIFFLWEKLPFNH 192

Query 189 AIWHLFVIAGTAFMYFCILFV 210
 AIWHLFV+ GA M+FC+LFV
 Sbjct 193 AIWHLFVLGSSAMMFCVLFV 214

Figure 2 - This is the first hit in the nr BLAST database for the gene. Its name is Hemolysin III and originates from the bacteria *Bacillus Cereus*. It's alignment length was 202 and it had an E-value of 9e-77

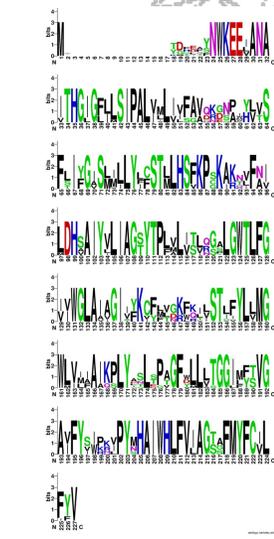


Figure 3 - This is a Weblogo, a graphical representation of an amino acid or nucleic acid multiple sequence alignment of the top BLAST hits. The large Green and Black letters represent Hydrophilic and Hydrophobic regions indicative of a transmembrane protein.

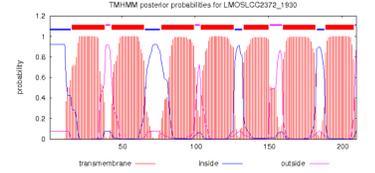


Figure 4 - This is a TMHMM transmembrane topology graph, which shows Hemolysin III has 7 Transmembrane helices.

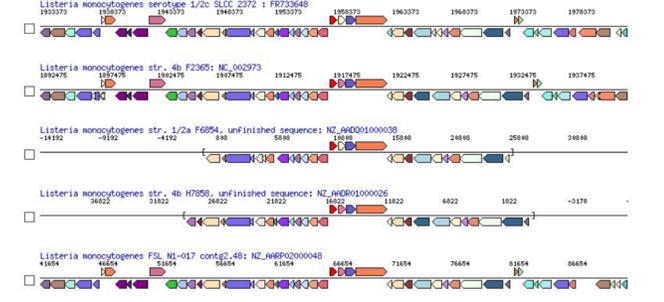


Figure 5 - This is an ortholog neighborhood region which shows that this gene is widespread in similar species and seem to be an integral part of *Listeria Monocytogenes*. It also indicates this gene is highly conserved through the species..



Figure 6-This is a Weblogo representing amino acid nucleic acid multiple sequence alignment of the top Blast hits. The large green and black letters represent hydrophilic and Hydrophobic regions of a transmembrane protein. Most of them are green(polar) and red(acidic)

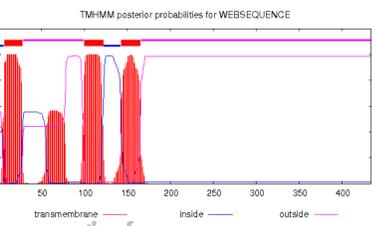


Figure 7-This is a TMHMM transmembrane topology graph which shows Hemolysin like CBS protein has 4 Transmembrane helices.

Figure 8- Hemolysin like CBS protein structure found on PDB

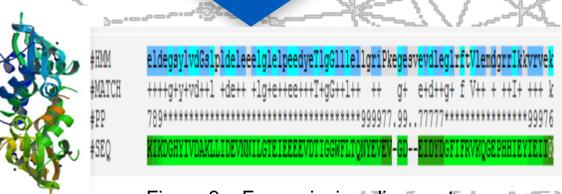


Figure 9: pFam pairwise alignment



Figure 10- KEGG Pathway Map of Hemolysin like CBS protein

Phylogenetic tree

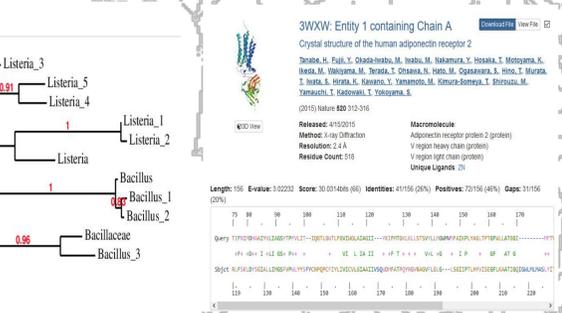


Figure 11- Phylogenetic Tree of *Listeria Monocytogenes*

Figure 12- Protein structure of Hemolysin III and pairwise alignment PDB

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.

Proposed Gene: Hemolysin III	Proposed Annotation: Hemolysin III
Proposed Gene: LMOSLCC2372_2300	Proposed Annotation: Hemolysin like CBS protein

References

Vázquez-Boland, J. A., Kuhn, M., Berche, P., Chakraborty, T., Dominguez-Bernal, G., Goebel, W., ... Kreft, J. (2001). *Listeria* Pathogenesis and Molecular Virulence Determinants. *Clinical Microbiology Reviews*, 14(3), 584-640. <http://doi.org/10.1128/CMR.14.3.584-640.2001>

Mickaël Desvaux and Michel Hébraud, Sept 2008
 Hemolysin supports survival but not entry of the intracellular bacterium *Listeria monocytogenes*. *M Kuhn, M. Kuhn, S. Kathariou, and W. Goebel, Jan 1988*

Acknowledgments

Supported by an NSF Innovative Technology Experiences for Students and Teachers (ITEST) Award - 1311902
 We'd Like to thank Dr. Steve Koury and Dr. Rama Dey-Rao for their support. www.buffalo.edu