Annotation of the *Listeria Monocytogenes* Genome with Locus Tags LMOSLCC2372_1000 and LMOSLCC2372_1001

Avaneesh Thoudojou, Shreyas Prasad, Sangeetha Gokhale
Williasmville East high School – Buffalo Engineering Awareness for Minorities and The Western New York Genetics in Research Partnership

Abstract

The successive grouping of the two genes derived from the organism *Listeria monocytogenes* were supported and analyzed in their structure, their phylogenetic bases, the degradation and duplication of possible gene sequences, possibility of horizontal gene transfer, cellular localization, potential alternative sequence and others within the gene annotation tool that is GENIACT. This is how the gene product name for the organism's protein was assessed, using the Genbank's varying subjects. The Genbank proposed product name for *Listeria monocytogenes* as the ABC transporter did not differ from the proposed gene annotation from the gene database. The reasoning why we chose these specific microbacteria to research was because, after extensive comparative analysis we uncovered that these were selected from the most common bacterium within cafeteria like environments in various food processing and handling utilities.

Introduction

Listeria monocytogenes is a facultative anaerobe, motile, non encapsulated, non-endospore forming a bacillus gram-positive bacteria, that is catalase positive and oxidase negative. It is responsible for causing food-borne illnesses, being a virulent foodborne pathogen. It was named after Joseph Lister and was described by E.G.D Murray first in 1924. It has become more prevalent over the past few years in the European Union. It causes listeric meningitis With symptoms including nausea, vomiting and diarrhea. The Anton test can be used in the identification *of L. monocytogenes*, while penicillin and ampicillin being possible treatments. Interestingly, Listeria Monocytogenes have been found in over 37 species of mammals, and many species of birds and fish. Listeria is also found in everyday food products like milk, cheeses, ice cream, raw vegetables, sausages and smoked fish.

According to some research done by Ramaswamy and others (2007), Immunocompromised individuals are more susceptible to *Listeria*, along with people on drugs for treating conditions such as cirrhosis, hemochromatosis and chronic renal, failure (those who require transfusions). What they also found was the process of pathogenesis within listeria, as it travels through the intestine, where the bacteria undergoes several divisions (Ramaswamy et al., 2007). These are usually controlled with a cell-mediated response from the body, but immunocompromised individuals are unable to respond and thus leading to low-level bacteremia and invasion of brain and gravid uterus. Some of their interesting findings include differences in invasion capacity from different kinds of cells between clinical and non clinical strains of *Listeria*, with Caco-2 and HepG2 cells showing differences in invasion capacity, showing that both these liver carcinoma cells are hugely affected by *Listeria* (Ramaswamy et al., 2007) . There are however some aspects that aren't known, including specific characteristics of the disease that Listeria Monocytogenes causes specifically, the timing for the onset of a serious form of Listeriosis and the timing for the gastrointestinal symptoms of it.

Our research of this gene from Listeria Monocytogenes is meant to reveal what the gene codes for, what are the characteristics of the protein, and what the protein functions as. We use tools such as WEB LOGO, BLAST, PFAM to find hits for our protein and display its biological properties, including chemical makeup and phylogeny. Our research also includes the location of the protein and served to find orthologs and paralogs, in addition to testing the protein for an RNA hit.

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

Methods

Modules of the GENI-ACT (http://www.geni-act.org/) were used to complete *Listeria monocytogenes* 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_1000:

The initial proposed product of this gene by GENI-ACT was an ABC Transporter ATP Binding protein. 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 transmembrane topography of the amino acid sequence and amino acid types as seen in WEB LOGO, and the cellular location of the amino acid sequence (being outside the cell, done using TMHMM). As such, the proposed annotation is an ABC Transporter ATP Binding protein.

LMOSLCC2372_1001:

The main gene expression product for this bacterium is ABC transporter permease supported through the gene analyzing tools in GENI-ACT. Based on the gene searching program that we used such as NCBI Protein Blast, where we compared the nucleotide sequence with the database they had stored on their gene data banks. The proposal that this was the gene product was shown through the top hits being ABC transporter permease protein located in the cytoplasmic region near the membrane boundary of the bacterium for more than the top 10+ hits with varying E values, but upper hits of the range having significant E values. There is also evidence through the cellular localization data discovered found using THMM, creating a WEBLOGO to demonstrate the structure that allows substrate movement in the cell, the transmembrane topography and helices indicating its role in the transport of molecules and other protein domain based functionalities.

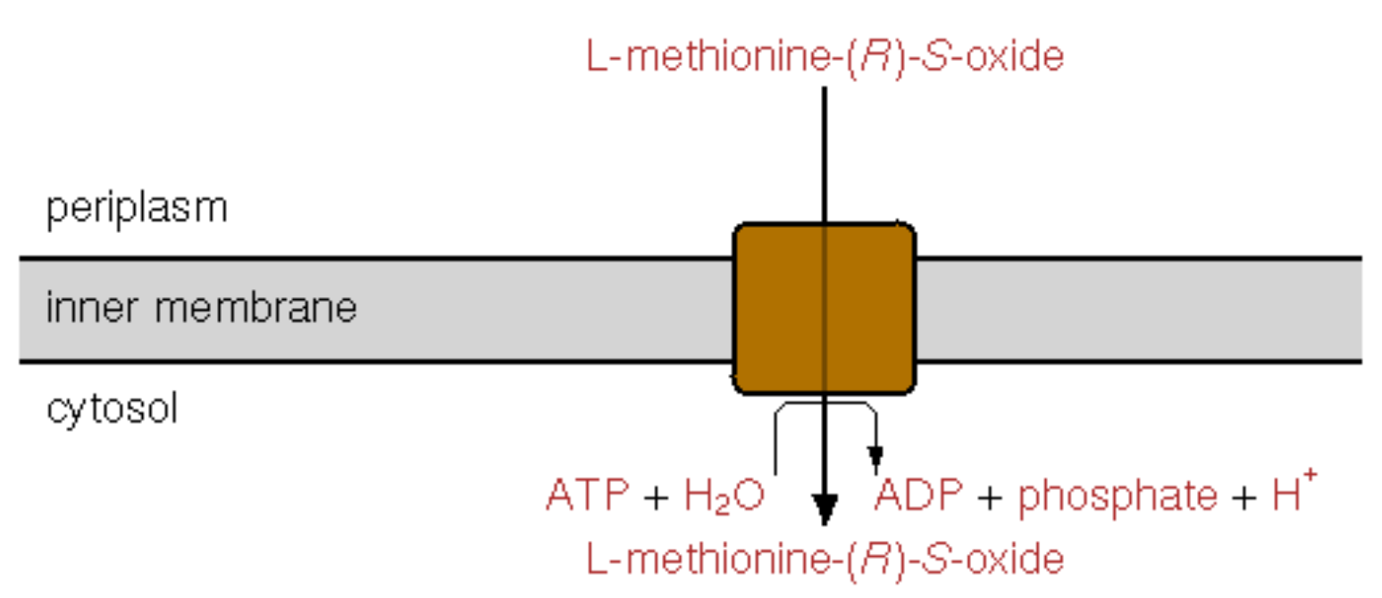
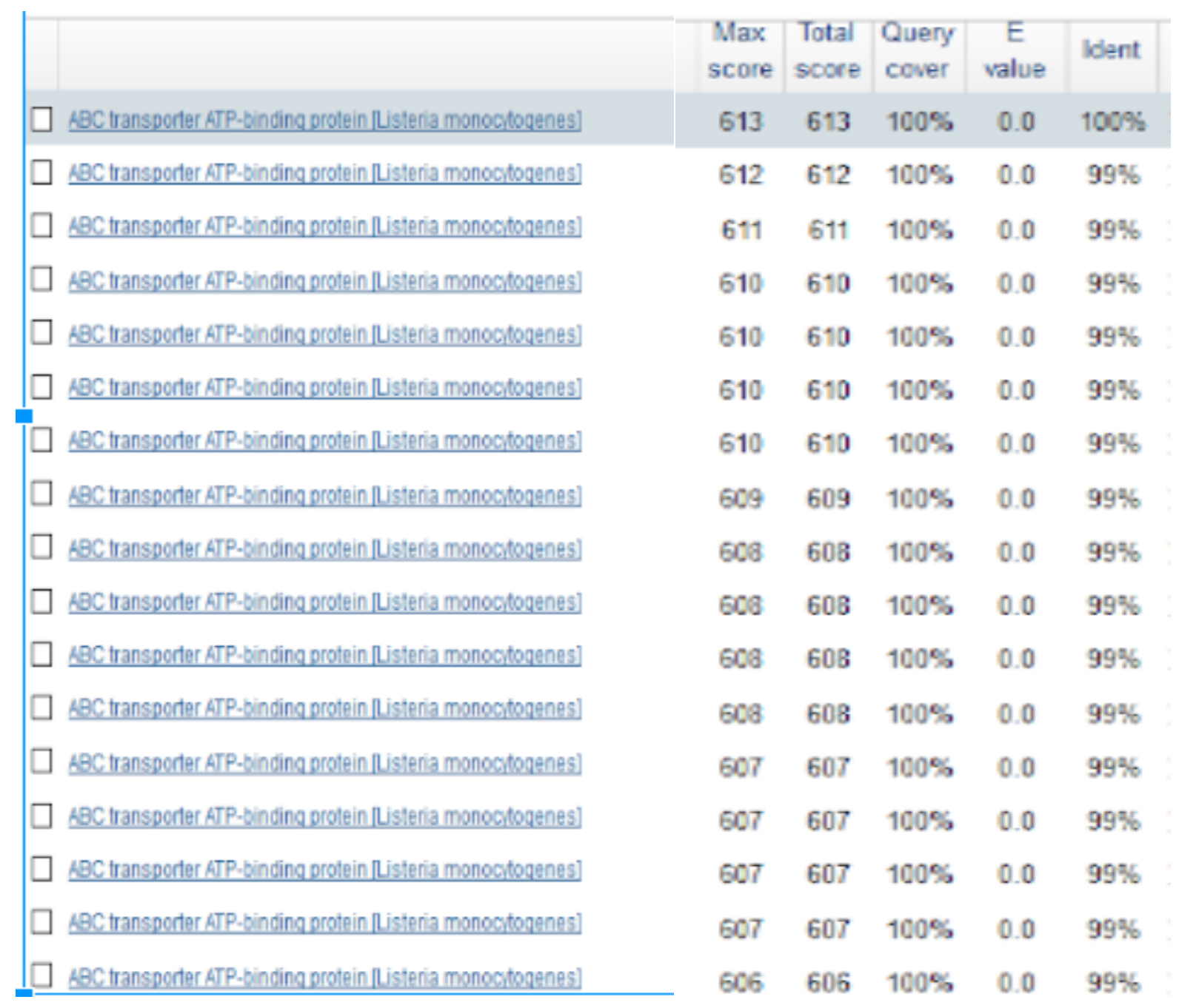


Figure 2. The direction of enzyme catalysis and shows that the ABC Transporter ATP Binding protein is involved in active transport (energized by phosphoanhydride bond hydrolysis)

Figure 3- The WEB LOGO for LMOSLCC2372_1000, with high occurrence of green indicating the high concentration of hydrophilic amino acids. This is congruent with the fact that an ATP Binding protein would need to have hydrophilic portions on it at sites of binding , since Adenosine Triphosphate (ATP) is also hydrophilic.



	Max score	Total score	Query cover	E value	Ident
☐ ABC transporter ATP-binding protein [Listeria monocytogenes]	613	613	100%	0.0	100%
☐ ABC transporter ATP-binding protein [Listeria monocytogenes]	612	612	100%	0.0	99%
☐ ABC transporter ATP-binding protein [Listeria monocytogenes]	611	611	100%	0.0	99%
☐ ABC transporter ATP-binding protein [Listeria monocytogenes]	610	610	100%	0.0	99%
☐ ABC transporter ATP-binding protein [Listeria monocytogenes]	610	610	100%	0.0	99%
☐ ABC transporter ATP-binding protein [Listeria monocytogenes]	610	610	100%	0.0	99%
☐ ABC transporter ATP-binding protein [Listeria monocytogenes]	610	610	100%	0.0	99%
☐ ABC transporter ATP-binding protein [Listeria monocytogenes]	609	609	100%	0.0	99%
☐ ABC transporter ATP-binding protein [Listeria monocytogenes]	608	608	100%	0.0	99%
☐ ABC transporter ATP-binding protein [Listeria monocytogenes]	608	608	100%	0.0	99%
☐ ABC transporter ATP-binding protein [Listeria monocytogenes]	608	608	100%	0.0	99%
☐ ABC transporter ATP-binding protein [Listeria monocytogenes]	608	608	100%	0.0	99%
☐ ABC transporter ATP-binding protein [Listeria monocytogenes]	607	607	100%	0.0	99%
☐ ABC transporter ATP-binding protein [Listeria monocytogenes]	607	607	100%	0.0	99%
☐ ABC transporter ATP-binding protein [Listeria monocytogenes]	607	607	100%	0.0	99%
☐ ABC transporter ATP-binding protein [Listeria monocytogenes]	607	607	100%	0.0	99%
☐ ABC transporter ATP-binding protein [Listeria monocytogenes]	606	606	100%	0.0	99%

Figure 4. Shows top hits from NCBI-BLAST, with all of the given hits showing an ABC Transporter ATP Binding Protein, with an E-Value of 0, a Query cover of 100% and an identity score of 100% or 99%

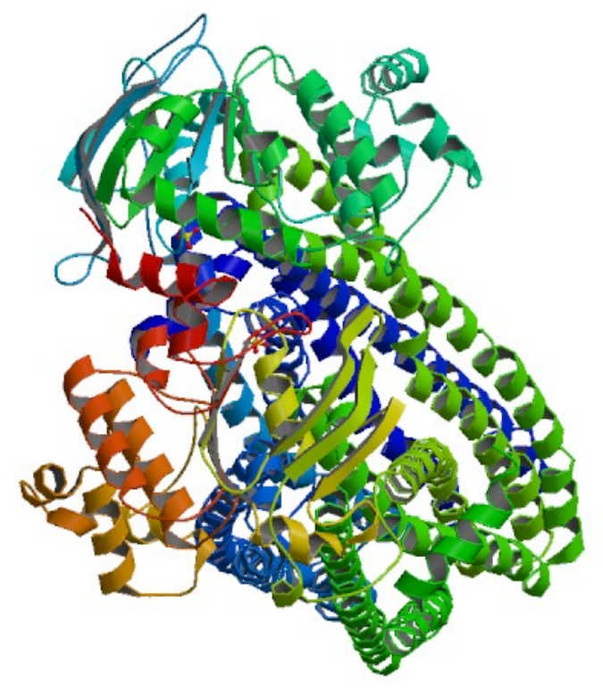


Figure 5 – Shows the ABC transporter permease structure, Its structure allows the substrates to pass through according to enzyme/protein regualtion and activation of the opening of the protein

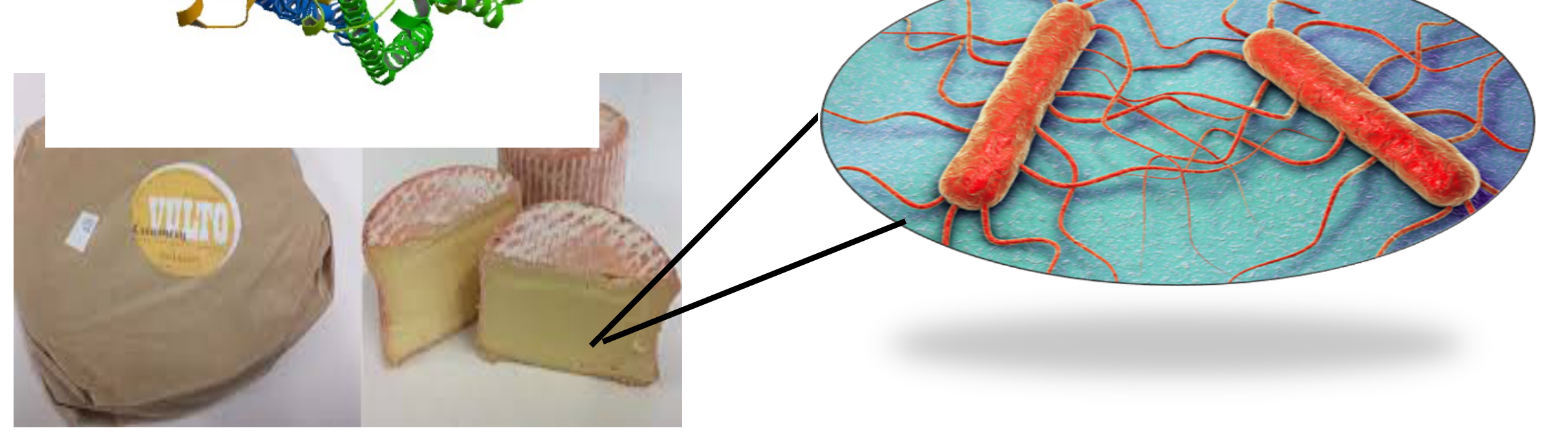


Figure 6 – This image shows the prevalence of Listeria Monocytogenes in our daily lives, as it has been found in cheese and has led to small outbreaks of Listeriosis

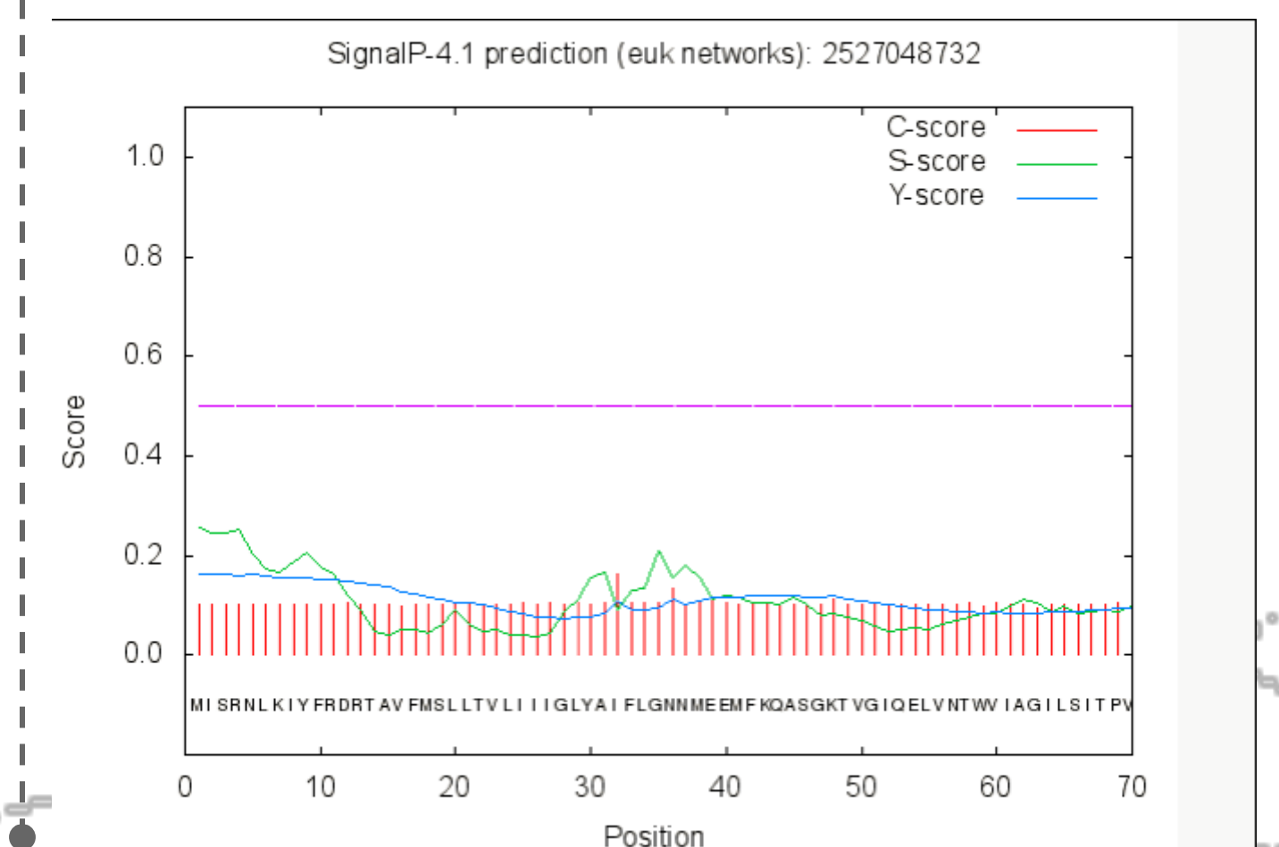


Figure 7: Demonstrated by the signal peptide graph we obtained from the SignalP protein source, we found that for ABC transporter permease,the C score (raw cleavage score) and the S score (signal peptide score) are both in a low range around 0.2 because of this and the low average value range of the Y score(combined cleavage site score) there are little to no chance of cleavage sites

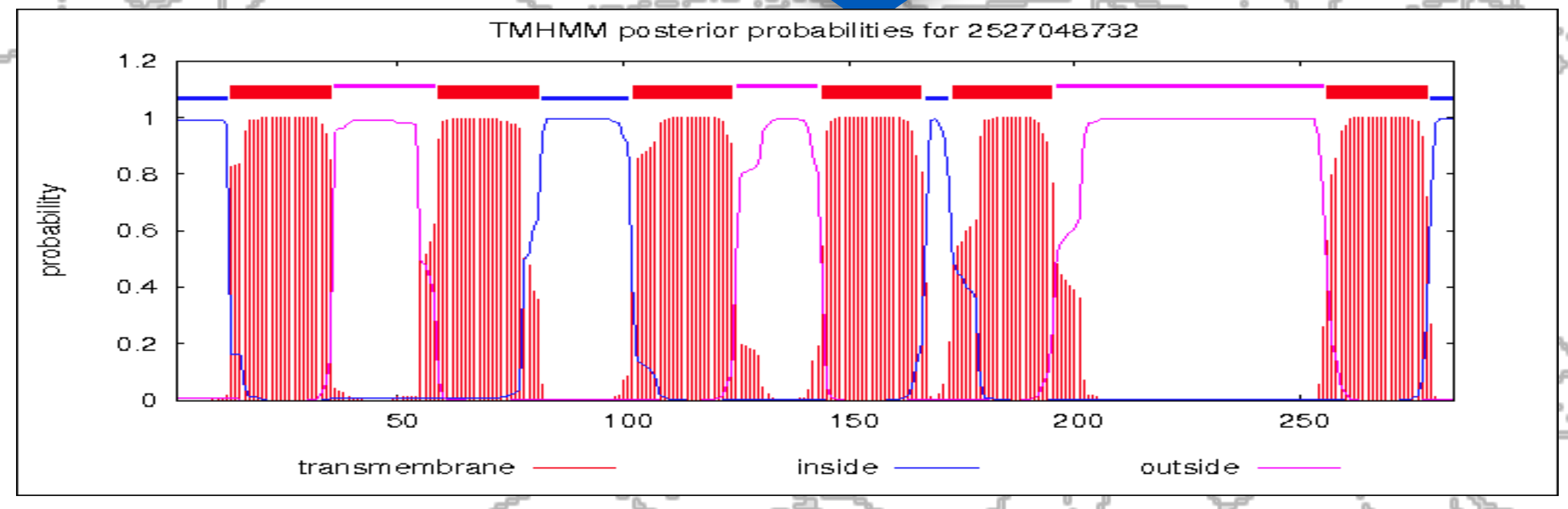


Figure 8: Found from the transmembrane topology rsult from TMHMM here it is depicted by the distinct red regions that there are 6 separate transmembrane helices, each of which are at or approaching the probability of 1 which means high chance that it is present

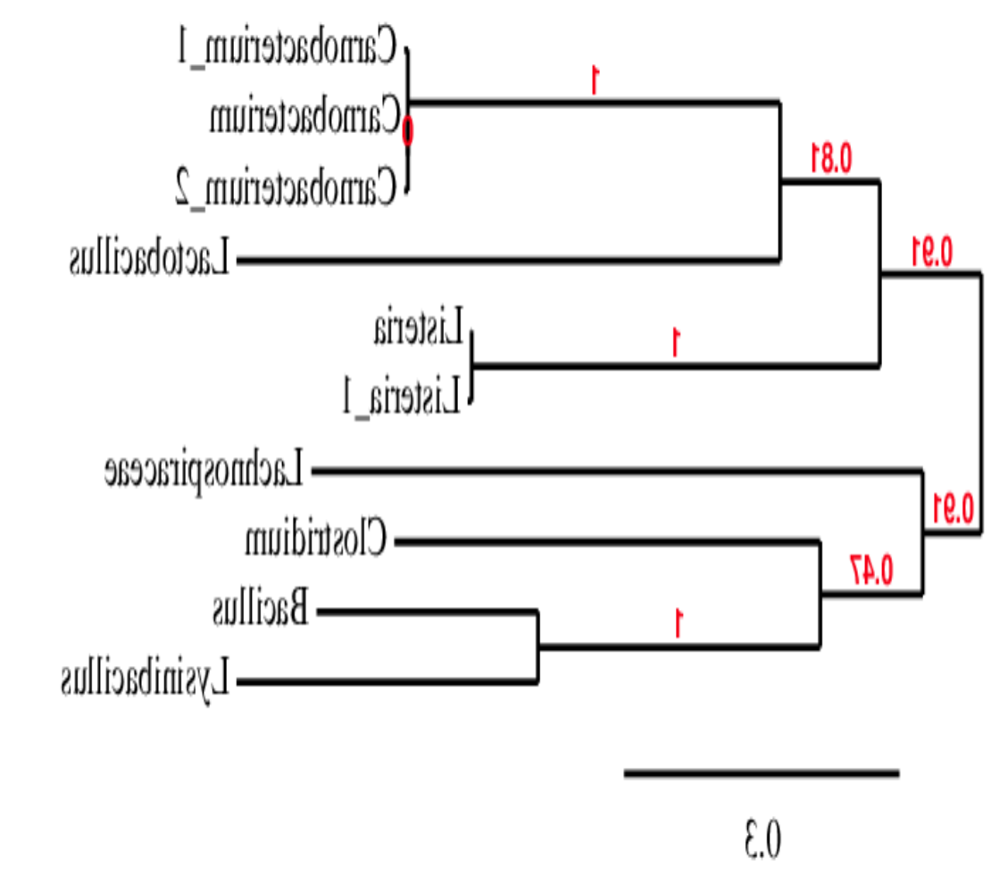
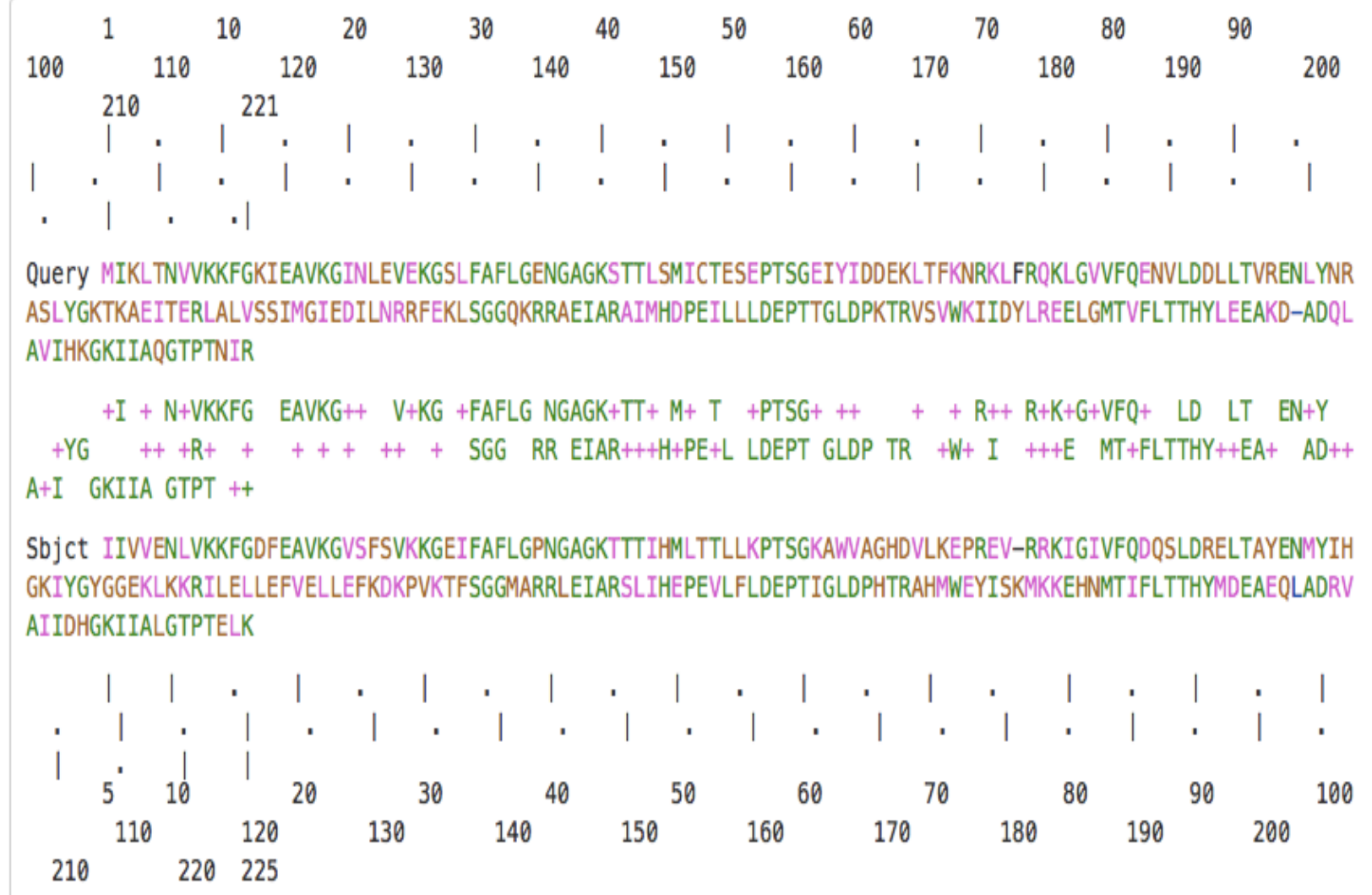


Figure 9- The given phylogenetic tree was made using the TReeDy tool on phylogeny.fr, showing the evolutionary relationships between different types of bacteria, with TreeDyn showing a close evolutionary relationship with Lactobacillus and Carnobacterium,

Figure 10: This image shows the pairwise alignment for the Crystal structure of an ABC transporter ATP-binding protein (TM_1403) from Thermotoga maritima MSB8 at 2.35 A resolution from PDB and a very low E-Value, at 1.86033E-52



```
Query MIKLTNVVKKFGKIEAVKGINLEVEKGSLFAFLGENGAGKSTTLSMICTESEPTSGEIYIDDEKLTFKNRKLFRQKLGVVFQENVLDDLLTVRENLYNR
ASLYGKTKAEITERLALVSSIMGIEDILNRRFEKLSGGQKRRAEIARAIMHDPEILLLDEPTTGLDPKTRVSVWKIIDYLREELGMTVFLTTHYLEEAKD-ADQL
AVIHKGKIIAQGTPTNIR

      +I + N+VKKFG  EAVKG++  V+KG +FAFLG NGAGK+TT+ M+ T  +PTSG+ ++     +  + R++ R+K+G+VFQ+  LD  LT  EN+Y
  +YG    ++ +R+  +   + + +  ++  +  SGG  RR EIAR+++H+PE+L LDEPT GLDP TR  +W+ I  +++E  MT+FLTTHY++EA+  AD++
A+I  GKIIA GTPT ++

Sbjct IIVVENLVKKFGDFEAVKGVSFSVKKGEIFAFLGPNGAGKTTTIHMLTTLLKPTSGKAWVAGHDVLKEPREV-RRKIGIVFQDQSLDRELTAYENMYIH
GKIYGYGGEKLKKRILELLEFVELLEFKDKPVKTFSGGMARRLEIARSLIHEPEVLFLDEPTIGLDPHTRAHMWEYISKMKKEHNMTIFLTTHYMDEAEQLADRV
AIIDHGKIIALGTPTELK
```

Conclusion

The gene product proposed by Geni-ACt and the database correlate with our findings that these genes code for an ABC Transporter ATP Binding Protein and an ABC Transporter Permease ,

Gene Locus	Geni-Act Gene Products	Proposed Annotatiom
LMOSLCC2372_1000	ABC Transporter ATP Binding Protein	ABC Transporter ATP Binding Protein
LMOSLCC2372_1001	ABC Transporter ATP Binding Protein	ABC Transporter ATP Binding Protein

References

Ramaswamy et al.(2007), Listeria— review of epidemiology and pathogenesis. Journal of Microbiology, Immunology and Infection 2007;40:4-13

Acknowledgments

Prof: Dr. Steve Koury and Dr. Rama Dey-Rao
University of Buffalo
www.buffalo.edu