Annotation of the *Clostridium botulinum* Genome at Locus Tags CLJ_0002, CLJ_0004 and CLJ_0279

Shaughnessy Shephard, Emily Frazier, Peyton Fontaine and Kirk Winters

Abstract

Three locus tags from the microorganism, *Clostridium botulinum* (CLJ_0002, CLJ_0004, CLJ_027) were annotated by 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, and cellular localization data of the gene. 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 there database.

Introduction

Clostridium botulinum is a gram-positive, anaerobe, endospore forming bacteria that is found is soil and untreated water throughout the world. The most distinguishable trait is *C. botulinum's* ability to produce the neurotoxin botulin. This neurotoxin is extremely deadly and is associated with causing flaccid muscular paralysis when injected as a medication called Botox. This neurotoxin can also be used to create Dysport, Xeomin, and Neurobloc.

There are several different types of botulism toxin classified by their antigenic specificity. The lethal dose for homo sapiens of this neurotoxin is 1ng/kg. This neurotoxin is also non-proteolytic, which means that affected food may look and smell normal. Also it is very hard to simply get rid of because to its heat resistant spores. Many canned foods which could contain this bacteria are pressured boiled which sufficiently kills the bacteria because of the extremely high temperatures. There are many different types of *C.botulism*. These include foodborne botulism, which is caused by the consumption of foods containing pre- formed botulinum toxin, wound botulism is cause by a toxin produced in a wound infected with *C.botulism*, infant botulism by definition occurs in persons less than one year of age and is caused by the consumption of spores of *C.botulism*, which then grow and release toxins in the intestines.

Figure 1:*Clostridium botulinum* is a rod-shaped bacterium and can be grown in a laboratory. The pH balance must be between 4.8 and 7.

Methods

Modules of the GENI-ACT (http://www.geni-act.org/) were used to complete *Clostridium botulinum* genome annotation. The modules are described in the chart 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?

Results

CLJ_0279

The initial proposed product of this gene by GENI-ACT was the protein helicase. This gene product proposal was supported by the first and second BLAST hits for the amino acid sequence. The first hit showed that it is a hypothetic protein which means its existence has been predicted. Sequencing of several genomes has resulted in prediction, but functions cannot be readily assigned. The second BLAST hit showed that it is punitive helicase; predicted to be helicase. The role of helicase is to unwind and separate the DNA strands during protein synthesis. As shown the proposed annotation is supported by the BLAST hits.

CLJ_0002

The initial proposed product of this gene by GENI-ACT was the protein Thermonuclease. This gene product proposal was supported by the BLAST hits for amino acid sequence. The first hit showed that it is the enzyme Thermonuclease which is Enzyme that catalyzes the hydrolysis of both DNA and RNA at the 5' position of the phosphodiester bond. The second BLAST hit showed that it is also Thermonuclease but the organisms have the same genus but are different species. The first is Staphylococcus intermedius which is gram-positive member of the bacterial genus Staphylococcus consisting of clustered cocci. The second hit is also Gram-positive member of the bacterial genus Staphylococcus consisting of clustered cocci.

CLJ_0004

The initial proposed product of this gene by GENI_ACT was a Botulism neurotoxin. The gene product proposal was supported by BLAST hits for protein sequences. The first hit showed that my neurotoxin is Botulinum neurotoxin type A. This specific neurotoxin is used in medicine to treat overactive muscle spams, this medicine is called BOTOX. My second BLAST hit also showed that my neurotoxin is Botulinum neurotoxin type A.

Data Analysis

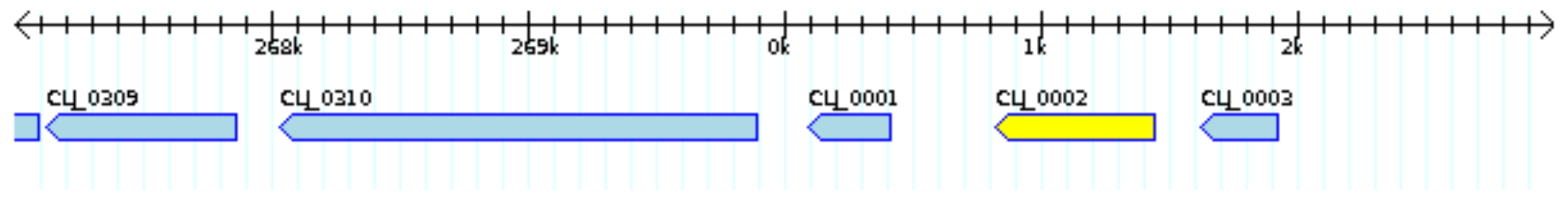

Figure 2. The gene neighborhood of CLJ_0002

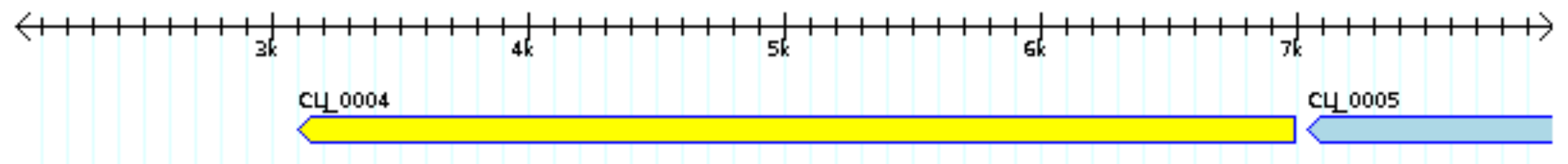

Figure 3. The gene neighborhood of CLJ_0004

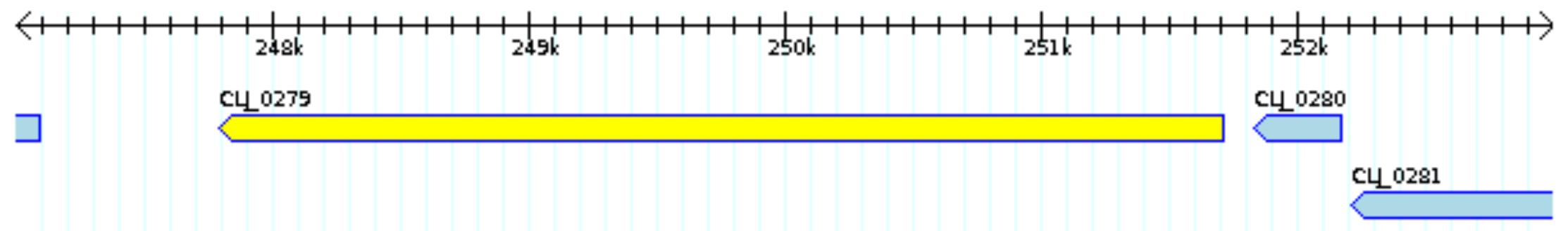

Figure 4. The gene neighborhood of CLJ_0279

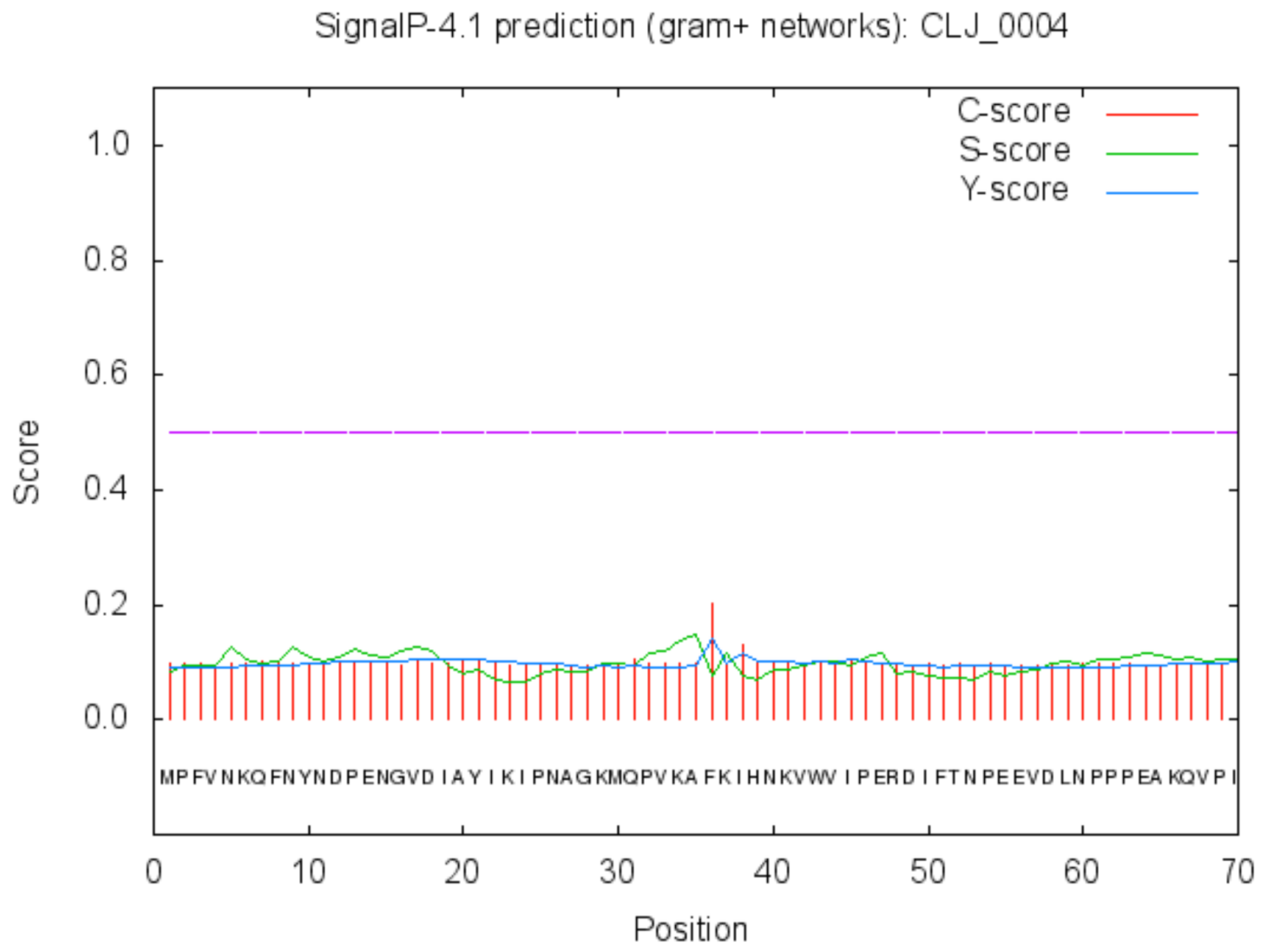

Figure 5: SignalP 4.1 Server- prediction results for CLJ_0004 showing NO signal peptide predicted

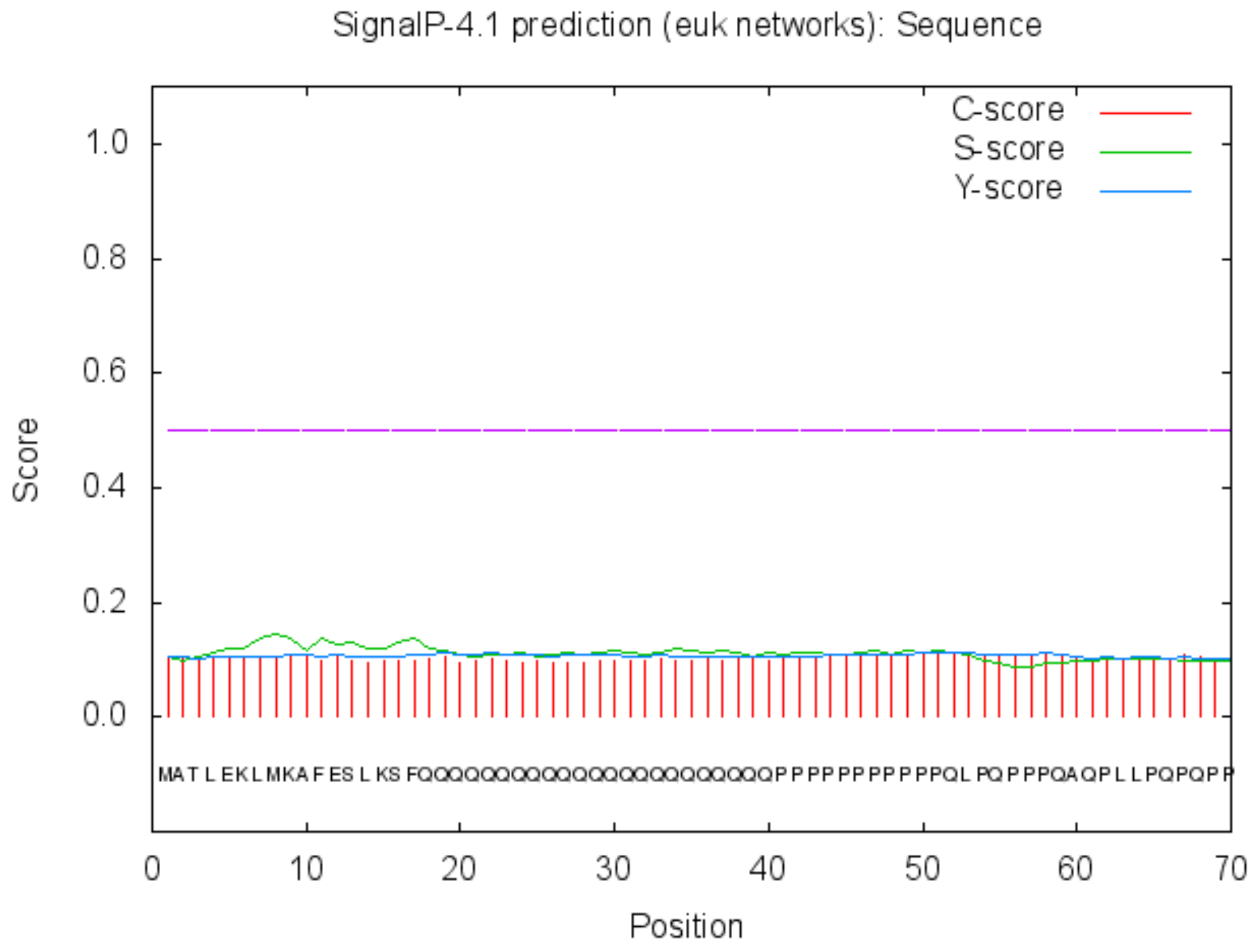

Figure 6: SignalP 4.1 Server Prediction results for CLJ_0279 saying that no signal peptide was found.

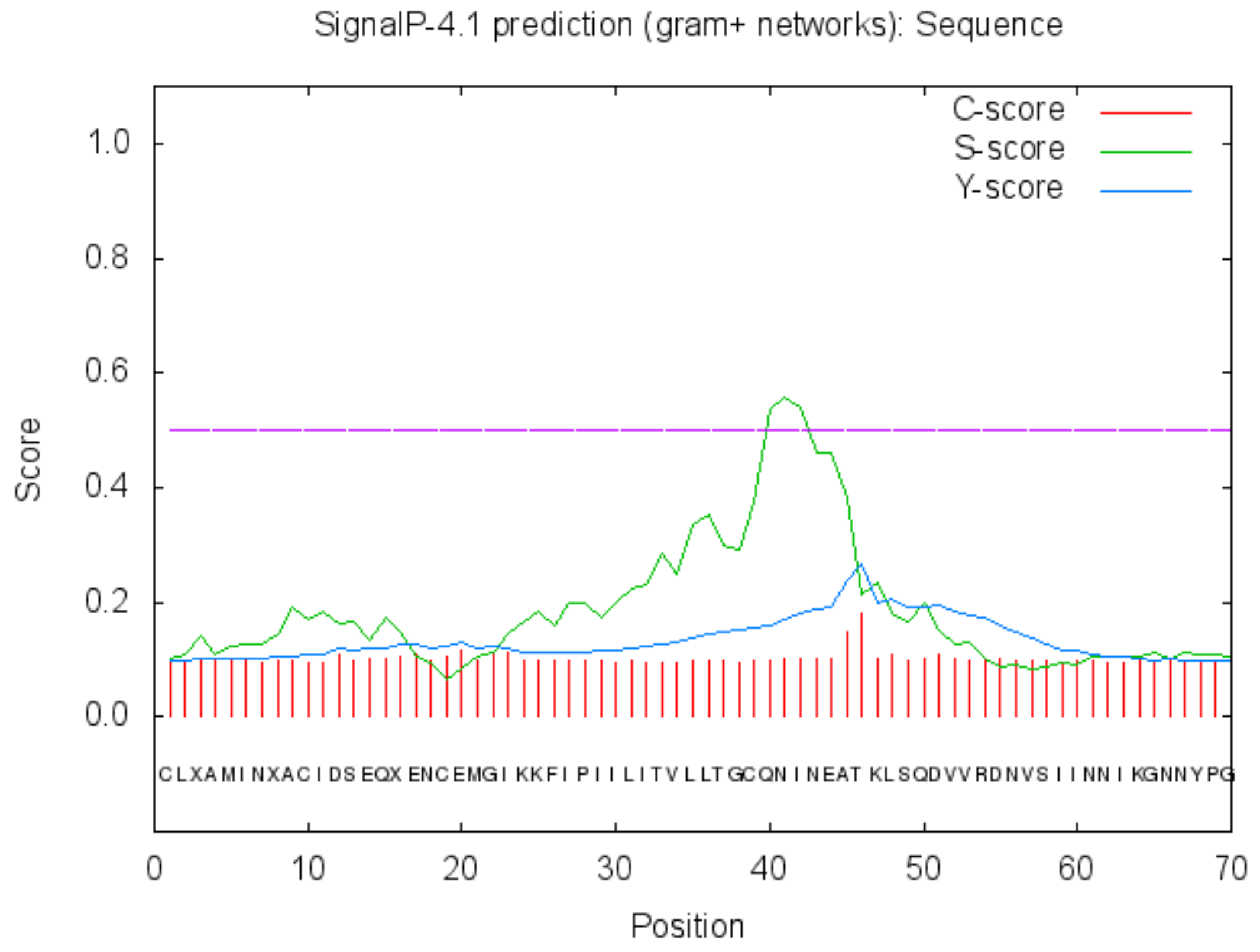

Figure 7: SignalP prediction results for CLJ_0002 predicting no signal peptide.

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.

GENE LOCUS	GENI-ACT GENE PRODUCTS	PROPOSED ANNOTATION
CLJ_0002	Staphylococcus	Protein Thermonuclease
CLJ_0004	Botulinum neurotoxin type A.	Botulism neurotoxin
CLJ_0279	Punitive helicase	Protein Helicase

Reference

https://web.mst.edu/~microbio/BIO221_2010/C_botulinum.html (for introduction)

Acknowledgements

Supported by an NSF Innovative Technology Experiences for Students and Teachers (ITEST) Award - 1311902

www.buffalo.edu