

Annotation of the *Yersinia pestis* Angola Genome at Locus

Tag YpAngola_0006

Mahum Awan, Mahjabin Ali, Na'Tazja Howell and Jennifer Livergood

Buffalo Academy of Science Charter School and the Western New York Genetics in Research and Health Care Partnership

Abstract

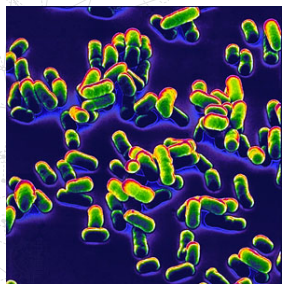
The gene located at Locus Tag YpAngola_0006 in the organism *Yersinia pestis* Angola (accession number CP000900) was annotated using the collaborative genome annotation website GENI-ACT. The Genbank proposed gene product name of pesticin immunity protein was assessed in terms of the general genomic information, amino acid sequence-based similarity data, structure-based evidence from the amino acid sequence, and cellular localization data. The two top BLAST hits with high scores and low e-values in nr database had the name of pesticin immunity protein. The only significant PFAM hit was PF16765 (PIM or Pesticin immunity protein). The Genbank proposed gene product name for YpAngola_0006 therefore seems correct.

Introduction

The Black Death to this day has been one of the most devastating pandemic in human history. For years, it elicited fear and death across parts of China and almost all of Europe. Back then, the Bubonic Plague was simply a mystery due to medicine and technology being under developed. It was short and simple: if you had it you would be dead in a couple hours but if you were lucky, in a few days. However, due to the advancement in medicine and scientific research, it has been established that a key cause of the Bubonic Plague is a bacteria called *Yersinia pestis*.

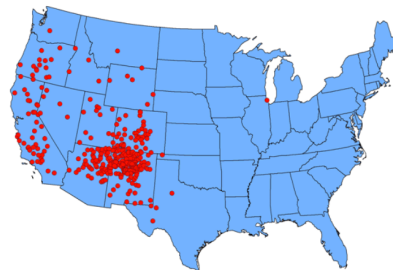
The genus *Yersinia* is composed of two other pathogens that cause infections in humans and animals: *Y. pseudotuberculosis* and *Y. enterocolitica*. The bacterium *Yersinia pestis* is transmitted to human beings by a flea bite of the dermis of an individual (Hu et al., 1998). Rodents and humans become the host organism for the bacterium, as it is passed from animal to animal. It can be passed aerosol, causing a pandemic. *Yersinia pestis* is a non-motile, gram negative bacterium (Eppinger et al., 2010).

Our study focuses on the gene located at Locus Tag YpAngola_0006.



Source: <https://www.britannica.com/science/Yersinia-pestis>

Reported cases of human plague—United States, 1970-2012



<https://www.cdc.gov/plague/maps/index.html>

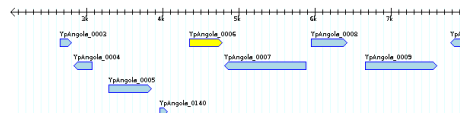


Figure 1. The locus tags and relative position of the genes under investigation in this research

Methods

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

Results

The proposed gene product name according to Geni-Act is pesticin immunity protein. The top BLAST hit was pesticin immunity protein. There were no COG numbers found for this gene. The only significant PFAM hit was PF16765 (PIM or Pesticin immunity protein). The T-Coffee alignment and weblogo showed a general lack of conservation throughout the sequence, but with some regions of conservation. PSORTb could not predict the cellular location of the protein. There were no predicted transmembrane helices according to TMHMM. The Phobius output confirmed this finding. However, both SignalP and Phobius predicted a signal peptide in the amino terminus of the protein, therefore suggesting it is secreted to the outside of the cell.

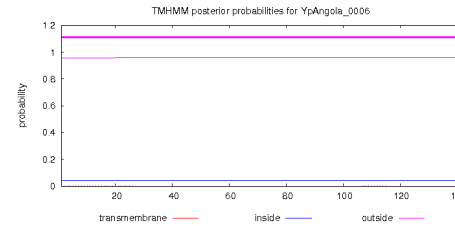


Figure 2. TMHMM showing no results for predicted transmembrane helices.

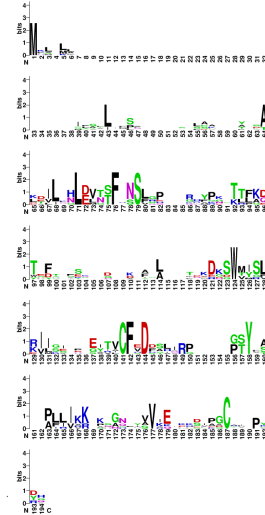


Figure 3. Weblogo results from YpAngola_0006. Little conservation is seen among the sequences used in the alignment, partially due to the fact that this protein seems fairly unique to *Yersinia pestis* (highly conserved among variants of *Yersinia pestis* in Genbank, but only a few other Genera with high scores and low e-values in BLAST

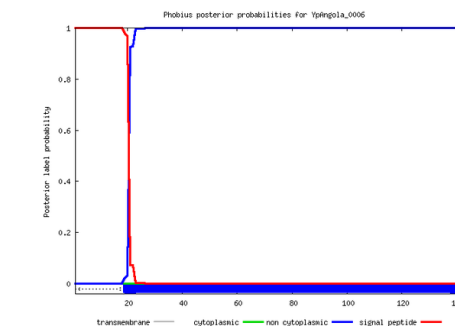


Figure 4. This graph represents the Phobius output, predicting a signal peptide, but no transmembrane helices.

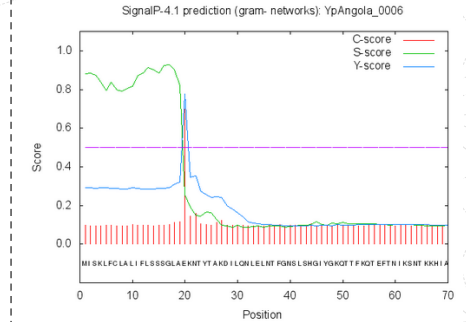


Figure 5. The SignalP output indicating that there is a signal peptide

Conclusion

The GENI-ACT proposed gene product name of pesticin immunity protein did not differ significantly from the gene annotation of YpAngola_0006 performed in this study. Pesticin is a toxin produced by *Yersinia pestis* that kills bacteria occupying the same environmental niche. Pesticin immunity protein is the immunity protein produced by *Yersinia pestis* and other Gammaproteobacteria to protect themselves against the bacteriostatic activity of the toxin (Patzner et al., 2012).

References

Chain et al. (2006). Complete genome sequence of *Yersinia pestis* strains Antiqua and Nepal516: Evidence of gene reduction in an emerging pathogen. *Journal of Bacteriology*, 188 (12), 4453-4463.

Eppinger, M., et al. (2010). Genome Sequence of the Deep-Rooted *Yersinia pestis* Strain Angola Reveals New Insights into the Evolution and Pangenome of the Plague Bacterium. *Journal of Bacteriology*, 192 (6), 1685-1699.

Hu, P., et al. (1998). Structural Organization of Virulence-Associated Plasmids of *Yersinia pestis*. *Journal of Bacteriology*, 180(19), 5192-5202.

Patzner SI, Albrecht R, Braun V, Zeth K; (2012). Structural and mechanistic studies of pesticin, a bacterial homolog of phage lysozymes. *J. Biol Chem*. 287:23381-23396..

Acknowledgments

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

www.buffalo.edu