Annotation of the *Propionibacterium acnes* Genome Locus Tags PAZ_c00320, PAZ_c00390, PAZ_c00430, PAZ_c00440

Saloni Chadha*, Angela Li*, Shikha Shelat*, Stephanie Terhaar*, and Mrs. Kathleen Kramer
Williamsville East High School and The Western New York Genetics in Research Partnership

Abstract

A group of four genes from the microorganism *Propionibacterium acnes* (PAZ_c00320, PAZ_c00390, PAZ_c00430, PAZ_c00440) 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, structure-based evidence from the amino acid sequence, cellular localization data, potential alternative open reading frames, enzymatic function, presence or absence of gene duplication and degradation, the possibility of horizontal gene transfer, and the production of an RNA product. 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 the database.

Introduction

Propionibacterium acnes is a Gram-positive human skin commensal that prefers anaerobic growth conditions and is involved in the pathogenesis of diseases. *P. acnes* bacteria live deep within follicles and pores and are known to grow in the lipid-rich environments of hair follicles. *P. acnes* may also be found within the gastrointestinal tract of humans and other animals. They are relatively slow-growing and are found to be an endophyte of plants, using the fatty acids in sebum and cellular debris as an energy source.

P. acnes is suspected to be discreetly involved in postoperative infections, prostheses failure, and more recently, in inflammation of lumbar nerve roots leading to sciatica. Previously known as *Corynebacterium parvum* in the early 1980s, it was used to stimulate the innate immune response against cancer in human and mice cells. P. acnes secretes many proteins, such as several digestive enzymes, aiding in the digestion of sebum and acquisition of other nutrients. They can also destabilize the layers of cells that form the walls of the follicle. It is thought that its growth may prompt inflammation inside the cells, evident in its major role with prostate inflammation. This inflammation, in turn, might spur the cells to become cancerous.
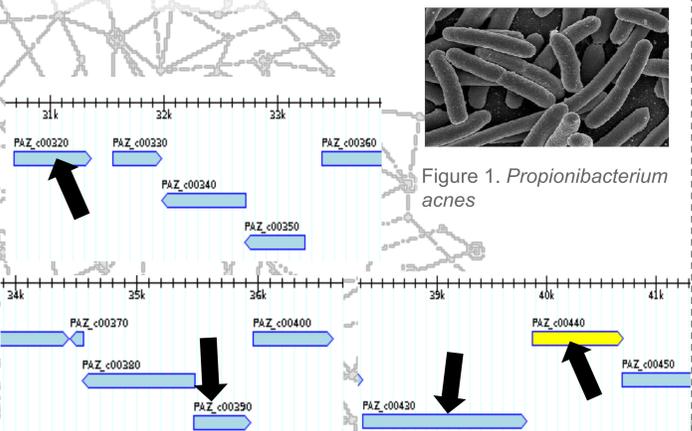

Figure 1. *Propionibacterium acnes*

Figure 2. 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 the Propionibacterium acnes 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

PAZ_c00320:

The initial proposed product of this gene by GENI-ACT was a transcriptional regulator. Transcriptional regulation is the means by which a cell oversees the conversion of DNA to RNA and controls how often a given gene is transcribed. This gene product proposal was supported by the HMM Logo results, pairwise alignments, curated alignments and SignalP results. The HMM logo specifies position-specific letter emission distributions and also position-specific insertion and deletion probabilities, describing the sequence family. The SignalP results indicate that no signal peptides are present, therefore implying that the initial proposed product is not a signal protein. As such, the proposed annotation is a transcriptional regulator.

PAZ_c00390:

The initial proposed product of this gene by GENI-ACT was a histone acetyltransferase. Acetyltransferases are enzymes that transfer an acetyl group from acetyl-CoA to lysine amino acids. Histone acetyltransferases can also acetylate non-histone proteins, to facilitate gene expression. The residues reside on histone tails and the WebLogo shows where the residues are reserved. The base pairing shown in the WebLogo are 3 Watson-Crick hydrogen bonds. As such, the proposed annotation is a histone acetyltransferase.

PAZ_c00430:

The initial proposed product of these genes by GENI-ACT was a Type II Secretion System Protein E. This gene product proposal was supported by the top BLAST hits for the amino acid sequences. A number of bacterial proteins,

some of which are involved in a general secretion pathway for the export of proteins (also called the type II pathway) belong to this group. These proteins are probably located in the cytoplasm and, on the basis of the presence of a conserved P-loop region, bind ATP. WebLogo results for PAZ_c00430 showed hydrophobic regions within the gene, causing amino acid links to fold inward within the gene.

PAZ_c00440:

The initial proposed product of this gene by GENI-ACT was a type II secretion system protein. This protein enables the transport of cytoplasmic proteins across the lipid bilayers that make up the cell membranes of the P. acnes organism. This proposal was supported by the transmembrane topography and the SignalP test results. The TMHMM results indicate the presence of three transmembrane helices, providing evidence that the gene product is indeed a protein that traverses the cell membrane and transports other proteins. Since the SignalP test showed that there are no signal peptides present, we can rule out that these transmembrane helices are signal proteins. Together, these data results show that the gene product is the type II secretion system protein.
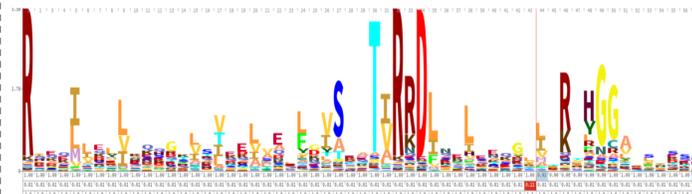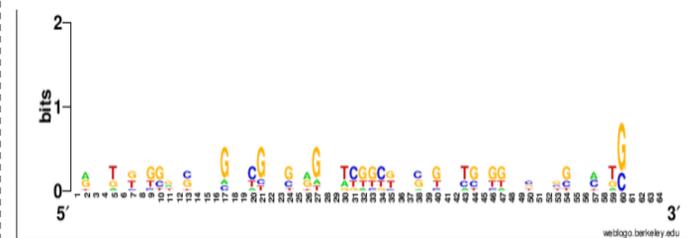

Figure 3- PAZ_c00320 HMM logo results.
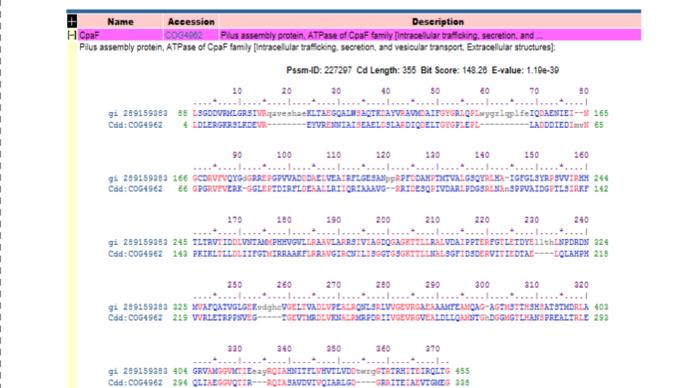

Figure 4 – PAZ_c00390 WebLogo results.

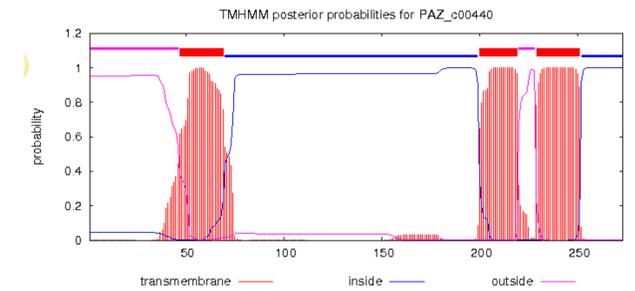Figure 5- PAZ_c00430 CDD search results, COG number top hit.
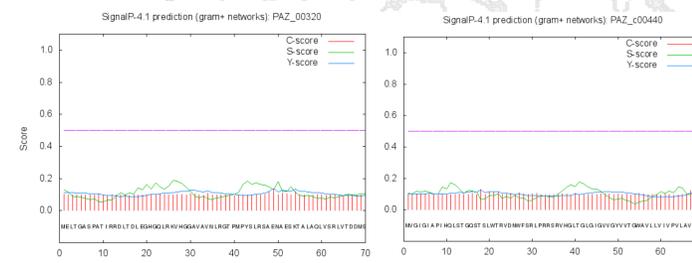

Figure 6 – PAZ_c00440 TMHMM results.

Figure 7- SignalP results for PAZ_c00320 (left) and PAZ_c00440 (right).

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 Tags	Geni-Act Gene Products	Proposed Annotation
00320	Transcriptional Regulator	Transcriptional Regulator
00390	Acetyltransferase	Acetyltransferase
00430	Type II Secretion System Protein E	Type II Secretion System Protein E
00440	Type II Secretion System Protein	Type II Secretion System Protein

References

Bhatia, Ajay. "PROPIONIBACTERIUM ACNES AND CHRONIC DISEASES." *The Infectious Etiology of Chronic Diseases: Defining the Relationship, Enhancing the Research, and Mitigating the Effects: Workshop Summary.* U.S. National Library of Medicine, 01 Jan. 1970. Web. 19 May 2017.

Acknowledgments

www.buffalo.edu