Isolation and Genomic Annotation of the Novel *Bacillus thuringiensis* Bacteriophage, Rex16

Sydney Dishman[1], Jennifer Easterwood[2], Kent Rhodes[2], Michael Wolyniak[3], Joanna Katsanos[2]

Affiliations: Georgetown University[1], Queens University of Charlotte[2], Hampden-Sydney College[3]

Contact Author : Sydney Dishman, sed75@georgetown.edu, (301) 997 4191

Citation : Dishman S, Easterwood J, Rhodes K, Wolyniak M, & Katsanos J. 2017. Isolation and genomic annotation of the novel *Bacillus thuringiensis* bacteriophage, Rex16. Q.U.E.S.T. Journal. 1: 19-25.

Keywords : Microbiology, Genetics, SEA PHAGES

ABSTRACT: Bacteriophages are exceptionally abundant, dynamic, ancient, and genetically diverse viruses that specifically infect prokaryotic organisms. With an estimated 10^{31} phages present on the planet at any given moment and less than 2,000 strains genetically characterized to date, bacteriophages represent a largely unexplored area of the global biome. Howard Hughes Medical Institute (HHMI) has sought to eliminate this gap in knowledge by establishing the Science Education Alliance Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program. It is through this program that the novel *Bacillus thuringiensis* bacteriophage, Rex16, was isolated and characterized. With a 162,605 base-pair genome, this strain was found to contain genetic information similar to other C1 subcluster phages as well as entirely unique open reading frames. The discovery of new bacteriophages, like Rex16, may progress the development of phage technology, which has shown promise in the fields of medicine and agriculture. In addition, this particular project has paved the way for continued bacteriophage isolation and genomic annotation at Queens University of Charlotte.

Introduction

Bacteriophages, parasitic viruses that specifically infect prokaryotes, were discovered simultaneously by Felix d'Herelle and Edward Twort in 1915. Like other viruses, phages are made up of genetic material (usually double-stranded DNA) surrounded by a protein shell called the capsid. Often, a tubular structure also made of protein, the tail, extends from the capsid for movement and attachment to cells. Phages proliferate by first attaching to the cell wall of a bacterium and injecting their own DNA into the host's cytoplasm. The viral genome is then replicated, transcribed, and translated into protein within the host cell. New phages are assembled from the replicated genetic material and translated phage proteins. As the amount of phage within the cell multiplies, the host begins to swell and eventually lyses from the pressure. The phages are released into the surrounding environment to locate and infect more bacterial hosts (HHMI, 2009).

Bacteriophages are exceptionally abundant, with 10^{31} individuals estimated to be present on earth at any given moment. They are also a dynamic, ancient, and genetically diverse group of organisms, containing many genes that have not previously been sequenced (Pope et al., 2015). With the introduction of the Science Education Alliance Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program by the Howard Hughes

Medical Institute (HHMI) in conjunction with the University of Pittsburgh in 2008, bacteriophage isolation and the sequencing of their genomes has exponentially accelerated. Almost 10,000 phages have been isolated and there are now nearly 2,000 completely sequenced phage genomes available for analysis (phagesdb.org). Despite the recent increase in novel phage isolation events, the number of sequenced genomes in the GenBank database represents an inconsequential fraction of the genetic information present in the biosphere. It has been established that multiple phage types specifically infect single bacterial species, indicating that bacteriophages are exponentially more abundant than their hosts (phagesdb.org). Although the ~2,000 recorded phage genomes reveal significant progress in this area of study, more than 30,000 prokaryotic genomes have been recorded in GenBank to date. This emphasizes the continued gap in knowledge and need for novel phage isolations and genomic annotations.

As more bacteriophages have been isolated and more genomic sequences have become available for analysis, inter- and intra-cluster relatedness has been revealed. The apparent mosaicism on the nucleotide and amino acid level indicate that phage diversity represents a continuum rather than distinct groupings (Pope et al. 2015). The mosaic property of bacteriophage genomes can be attributed to multiple factors. It has been proposed that phage proliferation follows Hendrix's pattern of inheritance. In this model, random genomic recombination occurs as the next generation of phage is produced. Although this creates numerous defective phage units and only a small fraction of functional individuals, it creates an opportunity for new and advantageous mutations to spontaneously appear within the phage genome. Bacteriophage genetic mosaicism can also be attributed to lateral gene transfer. In a study by Pedulla et al. (2003), 13 percent of open reading frames (ORFs) within a bacteriophage genome were not unique to that phage population. The majority of these ORFs were related to ORFs of their respective bacterial hosts. This supported the theory that genetic information is transferred between parasite and host during infection events (Papke, 2003).

Bacteriophage research is important because there is a potential for phage technology to be useful in industry. Their ability to inhibit the growth of bacteria while not affecting mammalian cells may be useful in developing disinfectants and antiseptics (McNerney, 1999). Phage technology has even been applied in the agriculture and food processing industry, with the first phage-based product, Agriphage™ produced by OmniLytics Inc., approved by the US Environmental Protection Agency in 2005. Additional products from Intralytix Inc., such ListShield™, EcoShield™, and SalmoFresh™ have been approved by the FDA for use in food processing and preparation by targeting *Listeria, E. coli,* and *Salmonella* cultures, respectively (Enderson, 2014). Since mycobacteriophages were discovered in 1947, their potential to be used in a medical environment has been discussed (Brodier-Montagutelli, 2016; Levin & Bull, 1996). Phage therapy faces multiple challenges, however. Within the body, bacteriophages are limited in their ability to locate and infect pathogenic cells. They also are likely to elicit an immune reaction that would both render the phage ineffective and has the potential to cause an unnecessary inflammatory response within an already-sick individual. In addition, phage-mediated lysis may result in the release of harmful byproducts which may lead to the development of secondary conditions such as sarcoidosis (McNerney, 1999; Zelasko, 2017).

The specificity of mycobacteriophages, however, may make them a useful tool in the diagnosing and treatment of bacterial infections. In an effort to prevent resistant strains of bacteria from proliferating, health care providers are reducing the use of general antibiotics in favor of more specialized alternatives. For this to be successful, however, the type of infectious bacteria must be identified as precisely as possible. It has been shown that phages have the ability to precisely identify strains of bacteria,

such as *Mycobacterium tuberculosis* (McNerney, 2005). Although this use of phage technology has not yet been applied beyond mycobacterial hosts, there is a potential for *Bacillus* bacteriophages to be used in the detection and treatment of infections caused by *Bacillus anthracis* and *Bacillus cereus*, the cause of anthrax and some food poisonings, respectively (Grose et al., 2014).

Bacteriophage isolation and identification can be used as an educational tool to introduce undergraduate students to authentic research. When done as a class and under the guidance of a professor, the process is relatively easy and engaging for students. Individuals who participate in this type of research garner important lab and writing skills, as well as develop a passion for microbiology and genetics. The Phage Hunters Integrating Research and Education (PHIRE) program includes eight steps that gradually increase in difficulty, challenging the students to progress from simply using concrete skills to later developing representational understanding. The research produces a genuine discovery for the scientific community, encourages peer-mentoring, has a high success rate, and is cost efficient (Hanauer, 2006).

It is through the SEA_PHAGES program that a novel *Bacillus thuringiensis* bacteriophage species, Rex16, was isolated from the soil at Queens University of Charlotte and characterized both phenotypically and genotypically.

Methods

Protocols were taken from the NGRI Bacillus Phage Laboratory Manual provided by the Science Education Alliance of the Howard Hughes Medical Institute and the Phagehunting Protocols posted to phagesdb.org by the University of Pittsburgh.

Phage Isolation

To use as a positive control, phage samples were collected from final purification spot plates provided by Hampden-Sydney College. From these samples, a sterilized high-titer lysate was prepared.

Soil was collected on March 24, 2015 near the Queens University of Charlotte sign in front of the Burwell building on Selwyn Avenue (Coordinates: 35.188356, -80.831594). The sample was enriched with Trypticase Soy (T-soy) Broth and *Bacillus thuringiensis* before being incubated at 30 °C and shaken at 180 rpm overnight. The enrichment mixture was then centrifuged and the supernatant was collected. To check for the presence of bacteriophages, a spot test was conducted. A mixture of this supernatant, *B. thuringiensis* and melted top agar (TA; 1X T-soy broth powder, 0.4% agar; 55 °C) was poured onto a T-Soy Agar plate. Separately, a T-soy agar plate was topped with TA and bacteria and divided into six equal sections. Enrichment dilutions of 10^{-1} to 10^{-4} as well as positive and negative controls were added to separate sections of this agar plate. The plates were incubated at 30 °C overnight.

A plaque from the 10^{-4} dilution grid of the spot test was gently scraped with a micro pipet tip and transferred to SM buffer (50 mM Tris-Cl, 8 mM $MgSO_4$-$7H_2O$, 100 mM NaCl, 0.01% gelatin, 1 mM $CaCl_2$). A streak plate was prepared by dividing a T-soy agar plate into three sections. The inoculated buffer was applied to one section with a wire loop, which was then flame-sterilized before spreading a small amount of the liquid to the next third. This was repeated with the final section of the agar plate before a mixture of TA and *B. thuringiensis* was layered on top of the streak plate. This protocol was repeated with sterile SM buffer (negative control) and control phage (positive control). The plates were incubated at 30 °C overnight.

Three plaques from the most dilute third of the streak plate, labeled A, B, and C, were collected using micro pipet tips and transferred to separate aliquots of SM buffer. Serial dilutions 10^{-1} to 10^{-4} were prepared from each of the three samples. Each dilution was added to a mixture of melted TA and *B. thuringiensis* before being

poured onto separate T-soy agar plates. Positive and negative control plates were also prepared. All fourteen plates were incubated overnight at 30 °C. The most dilute plate containing clearly distinct plaques was chosen and three plaques from this plate were selected and labeled. This titer protocol was repeated four additional times, using three plaques and 10^{-1} to 10^{-4} serial dilutions each round (Figure 1).

After the fifth titer, plaque morphology was recorded and plates were checked for any indication of more than one phage population. A medium titer lysate was prepared by flooding a 10^{-2} dilution plate displaying "web pattern lysis" with SM buffer. The saturated top agar was scraped, collected and centrifuged. The supernatant was filter sterilized and labeled "medium titer lysate."

A single T-soy agar plate was divided into twelve sections. A mixture of melted TA and *B. thuringiensis* was poured onto the plate and allowed to solidify. Serial dilutions 10^{-1} to 10^{-10} of the medium titer lysate were prepared and added to ten separate grids on the agar plate. Control phage lysate and SM buffer were used as positive and negative controls, respectively. After overnight incubation at 30 °C, it was determined that a dilution of 10^{-4} produced "web pattern lysis." A 10^{-4} dilution of the medium titer lysate was added to melted TA and *B. thuringiensis* and this mixture was added to five separate T-soy agar plates. After solidification, the plates were incubated overnight at 30 °C with positive and negative control plates.

A high titer lysate was produced by flooding each of the five "web pattern lysis" plates with SM buffer and collecting the resulting slurry into a single conical tube. The mixture was centrifuged and the supernatant was collected and filter sterilized.

DNA Purification

Residual bacterial DNA within the high titer lysate was first degraded with a nuclease mix (150 mM NaCl, 0.25 mg/mL DNase I, 0.25

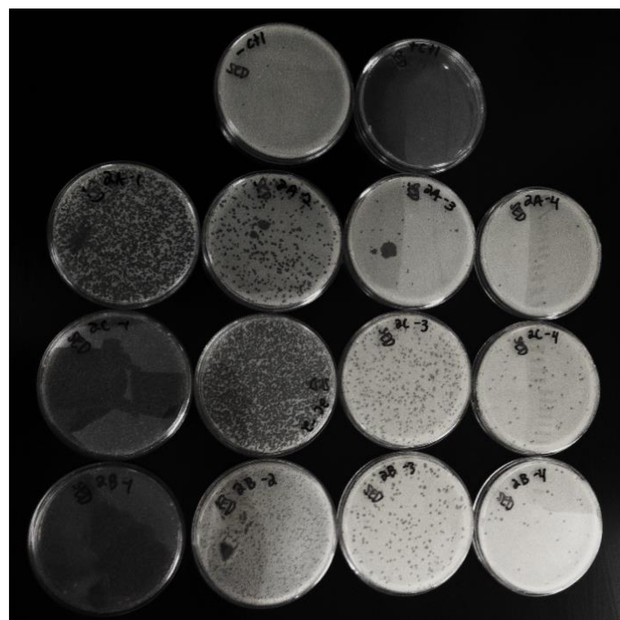


Figure 1 - Example of a three-phage, four-fold dilution titer. Top left: positive control; Top right: negative control; Four columns left to right: 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1} dilutions; Three rows top to bottom: Phage A, B, and C.

mg/mL RNase A, 50% glycerol). Phage particles were then precipitated by adding a phage precipitant solution (30% PEG 8000, 3.3 M NaCl) to the lysate. The mixture was centrifuged and the supernatant was discarded.

Phage genomic DNA was then isolated using the Promega Clean-Up Kit. The pellet was re-suspended in DNA Clean-Up Resin and the DNA became bound to a membrane as the solution was pushed through a column with a syringe. Isopropyl alcohol (80%) was then pushed through the membrane to wash away salts and proteins. The membrane was dried completely through centrifugation. The purified DNA was then collected by passing pre-warmed sterile water through the membrane into a eppendorf tube using centrifugation.

Nanodrop technology was used to analyze the sample of phage DNA. The concentration was calculated at 283.8 ng/µl with a 260nm/280nm ratio of 1.93.

Restriction Digest

Samples of phage genomic DNA were digested by five enzymes (BamHI, ClaI, EcoRI, HaeIII, and HindIII) separately in 10X reaction buffer and 10X bovine serum albumin (BSA). Loading dye was added to these solutions as well as a solution containing undigested phage DNA. A 0.8% agarose gel was prepared with GelRed fluorescent DNA stain and transferred to an electrophoresis chamber filled with 1X Tris/borate/EDTA (TBE) buffer. The gel was loaded with a 1 kb Ladder, the undigested DNA sample, and the five digested DNA samples. The gel was allowed to electrophorese at 75V for approximately one hour. The gel was then removed and photographed under UV light (Figure 2).

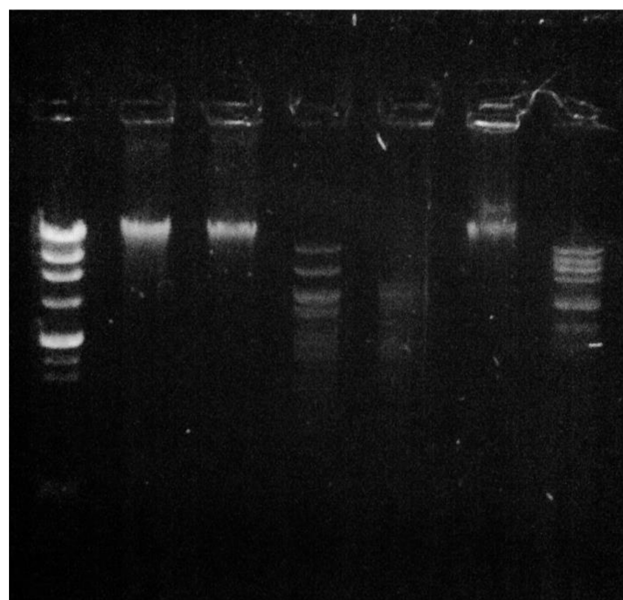


Figure 2 - Restriction digest on extracted phage DNA. From left to right: 1 kb ladder, undigested DNA, BamHI, ClaI, EcoRI, HaeIII, and HindIII.

Electron Micrograph

Highly concentrated samples of phage in both SM buffer and phosphate-buffered saline (PBS) were sent to the Electron Microscopy Lab at the Cannon Research Center in Charlotte, NC (Carolinas Healthcare System). The samples were prepared using the Drop-to-Drop Method and digital images were taken using a transmission electron microscope. Capsid size and tail length were estimated from these micrographs.

Genome Sequencing and Finishing

Isolated genomic DNA was sequenced by the third-generation sequencing technology, Illumina, at the Cannon Research Center. A total of 440,814 short reads were generated in a .fastq format.

The SEA-PHAGES Virtual Machine provided by HHMI and the University of Pittsburgh was used to assemble Rex16's genome. The complete .fastq file was initially used by GS De Novo Assembler for this task, but 41 different contigs ranging from 42,495 bp to 100 bp were produced. To solve this issue, the .fastq file was down-sampled to contain only the first 50,000 reads. A second run of GS De Novo Assembler using this new file resulted in a complete consensus sequence of 162,605 bp. The sequence was checked for low coverage and signaling conflict, although no areas of concern were found. The .fasta file is available on phagesdb.org.

Genome Annotation

Rex16's entire consensus sequence was imported into DNA Master, a freeware program designed by Dr. Jeffery Lawrence of the University of Pittsburgh. The programs Glimmer, GeneMark, and Aragorn were used in an auto-annotation of the genome. Each of these proposed open reading frames (ORFs) were compared to others currently recorded in GenBank using the National Center for Biotechnology Information's (NCBI) nucleotide Basic Local Alignment Search Tool (BLASTn). DNA Master was also used to produce a six-frame translation map and provisional genome map. A graph of coding potential for these ORFs was produced using *B. thuringiensis* as a model using GeneMarkS version 2.8. Although the outputs of the Phamerator and Starterator programs included in the SEA-PHAGES Virtual Machine are also recommended for the annotation process, lack of administrative access

to the Virtual Machine kept these visuals from being produced.

Using the quantitative evidence produced by the forenamed programs, along with the Guiding Principles of Bacteriophage Genome Annotation as outlined by SEA, Rex16's automatically produced draft annotation was evaluated and improved. Abnormally short ORFs and those weakly supported by available evidence were deleted while genes were added to particularly large gaps in the sequence. Possible alternative start sites were evaluated and some gene lengths were altered to include all protein coding potential. The updated annotation was re-BLASTed and a final genome map was produced.

Results

Electron Micrograph

Data from electron microscopy indicated that Rex16's morphotype falls under the Myoviridae category. Capsid diameter measurements ranged from 76.3 nm to 98.0 nm, with an average of 85.0 nm. Tail length ranged from 83.6 nm to 127 nm, with an average of 116nm (Figure 3).

Genetic Sequence

The complete genome of Rex16 is made up of 162,605 base pairs. The percentage of guanine and cytosine nucleotide bases is 38.8%. The BLASTn analysis of the entire genome indicates that Rex16 has a 98% identity score with the *Bacillus* bacteriophage Hakuna, and 95% identity scores with the *Bacillus* phages Megatron and Eyuki. These phages all fall into the subcluster C1, providing strong support for Rex16's identity as a member of the C1 subcluster as well. The next most similar hits include the *Bacillus* phages BPS13, BPS10C, and W.Ph. with identity scores of 86%, 85%, and 83%, respectively (Figure 4).

Genome Annotation

The genome of Rex16 was subdivided into 299 ORFs. The majority of these genes and predicted protein products were significantly (E-Value of 0.0E0) related to those of Bacillus phages Hakuna, Megatron, Eyuki, BPS13, BPS10C, and W.Ph. (Figure 5). While the function of some gene products could be determined because of their similarity to previously analyzed proteins recorded in GenBank, a majority of ORFs had to be labeled as hypothetical genes. Full annotation notes available in a .dnam5 file.

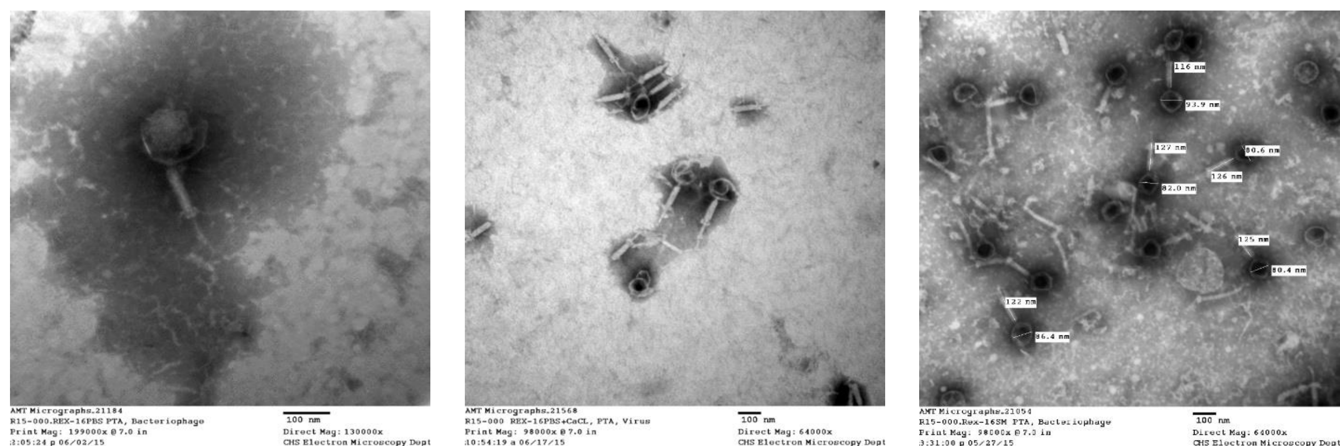
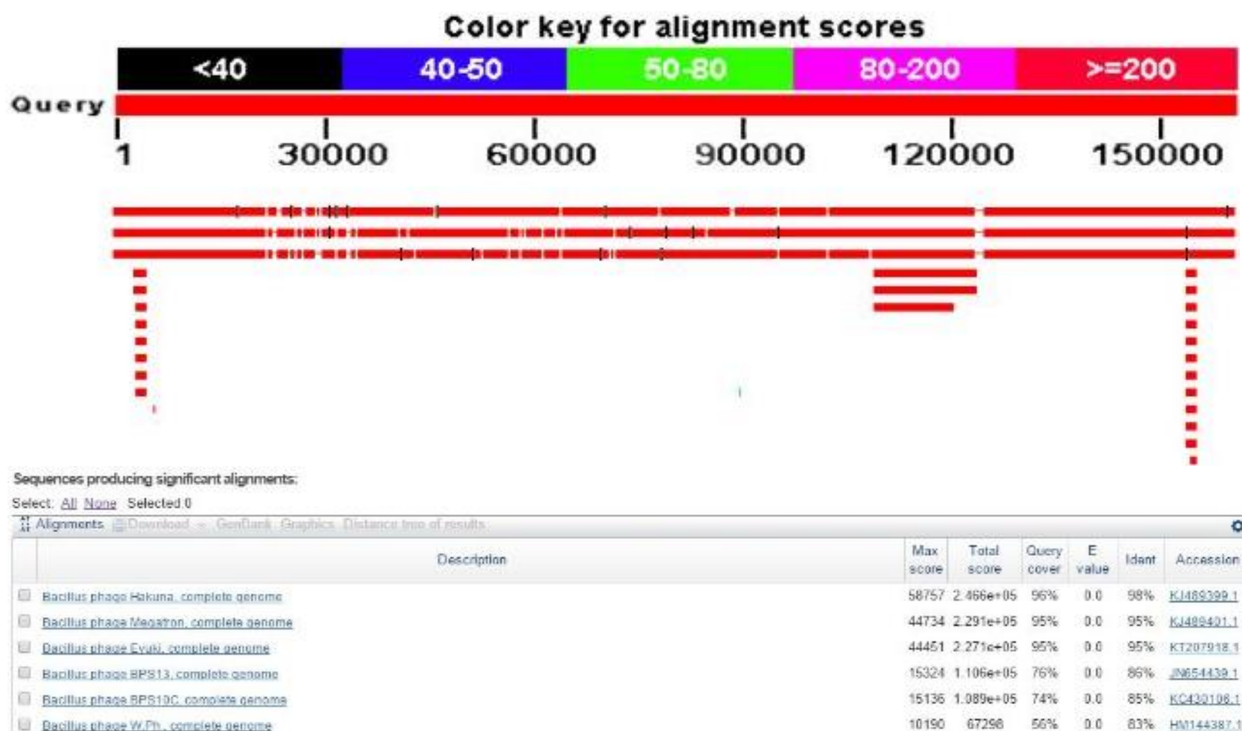


Figure 3 - Electron micrographs of Rex16. Magnification from left to right: 130,000x, 64,000x,



Description	Max score	Total score	Query cover	E value	Ident	Accession
Bacillus phage Hakuna, complete genome	58757	2.466e+05	96%	0.0	98%	KJ489399.1
Bacillus phage Megatron, complete genome	44734	2.291e+05	95%	0.0	95%	KJ489401.1
Bacillus phage Eyuki, complete genome	44451	2.271e+05	95%	0.0	95%	KT207918.1
Bacillus phage BPS13, complete genome	15324	1.106e+05	76%	0.0	86%	JN654439.1
Bacillus phage BPS10C, complete genome	15136	1.089e+05	74%	0.0	85%	KC430106.1
Bacillus phage W.Ph., complete genome	10190	67298	56%	0.0	83%	HM144387.1

Figure 4 - Screenshot of BLASTn results for full genome of Rex16. Top results included the Bacillus phages Hakuna, Megatron, Eyuki, BPS13, BPS10C, and W.Ph.

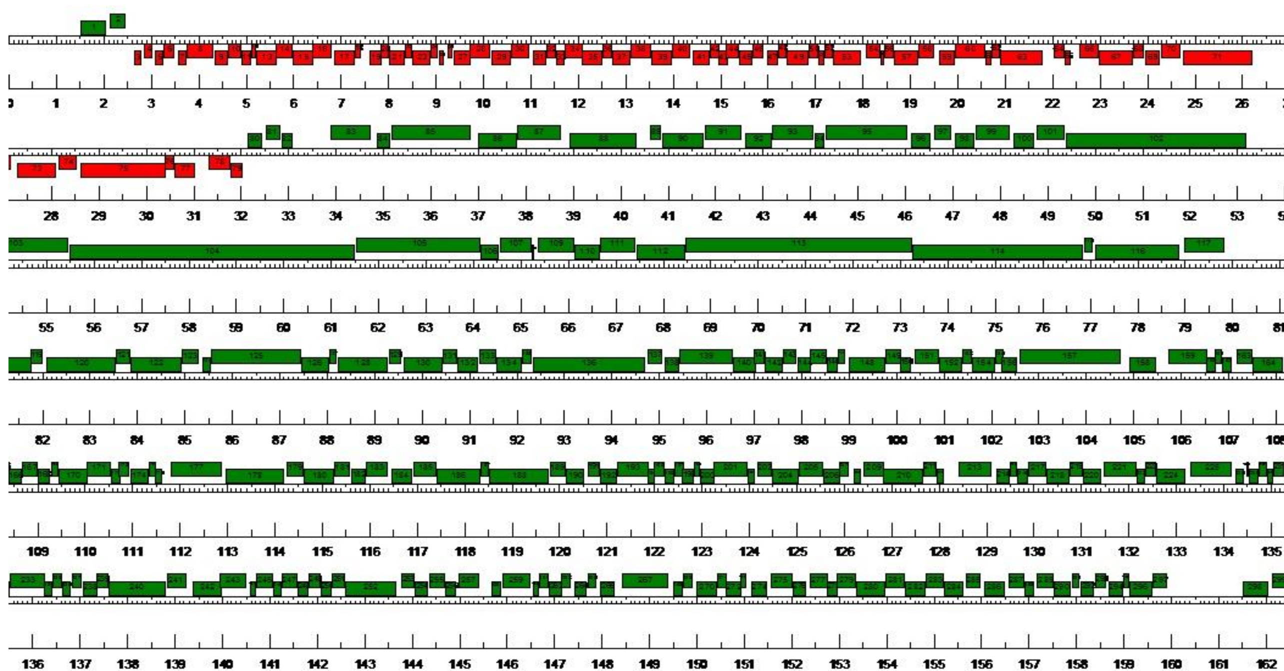


*Figure 5 - Genome map of Rex16. Numbers within the boxes indicate the ORF number (299 total) while the numbers along the line indicate the base pair location of the ORFs in thousands of bp (*162,605 *total). Green represents forward-transcribed ORFs and red represents reverse-transcribed ORFs.*

Discussion

The successful isolation and annotation of the *Bacillus thuringiensis* bacteriophage, Rex16, has furthered our understanding of the exceptionally abundant, dynamic, ancient, and genetically diverse phage population. While very similar to other cluster C1 bacteriophages, Rex16's unique genetic characteristics support the theory of nucleotide mosaicism in the phage population and the existence of a diversity continuum (Pope et al. 2015). These characteristics are likely due to the horizontal transfer of genetic material and Hedrix's theory of inheritance (Pedulla, 2003).

While the use of *Bacillus* bacteriophages has yet to be explored, it is assumed that they could be as successful in industry as their phage relatives. Phage technology has shown promise in its ability to create environmentally friendly disinfectants and antiseptics, to rid food products of unwanted pathogens, to diagnose and treat medical conditions, and to be used in the transfer of vectors in genomic research (Brodier-Montagutelli, 2016; Enderson, 2014; McNerney, 1999; McNerney, 2005). With additional research and product development, *Bacillus* bacteriophages' unique ability to specifically infect *Bacillus* hosts could be applied to all of these industries. The pathogens *B. anthracis* and *B.cereus*, for example, could be targeted by *Bacillus* phage technology, eliminating anthrax infections and some food poisonings.

The SEA-PHAGES program has benefited science students around the country. The integration of real research into the undergraduate classroom setting has not only provided students with the opportunity to develop important technical skills in microbiology and genetics, but has also increased their interest in and appreciation for biological studies (Hanauer, 2006). Through this process, I have provided Queens University of Charlotte with a model that the biology department will follow while integrating the SEA-PHAGES program into undergraduate courses in fundamental biology and genetics.

Acknowledgements

We would like to acknowledge Daisy M. Ridings and Ben J. Wagner of the Electron Microscopy Laboratory at Carolinas HealthCare System for the negative staining images obtained with the Philips CM10 Transmission Electron Microscope, as well as Dr. Nury Steuerwald and the team at the Molecular Biology and Genomics Core Facility at the Levine Cancer Institute in Carolinas Medical Center for the genomic sequencing of Rex16.

References

Brodier-Montagutelli E, et al. 2016. Inhaled phage therapy: A promising and challenging approach to treat bacterial respiratory infections. Expert Opinion on Drug Delivery. 10: 1-14.

Enderson L, et al. 2014. Phage Therapy in the Food Industry. Annual Review of Food Science and Technology. 5: 327-49.

Grose JH, et al. 2014. Genomic comparison of 93 Bacillus phages reveals 12 clusters, 14 singletons and remarkable diversity. Genomics. 15(855): 1471-2164.

Hanauer DI, et al. 2006. Inquiry learning. Teaching scientific inquiry. Science. 314(5807): 1880-81.

Howard Hughes Medical Institute (HHMI). 2009. NGRI Phage Laboratory Manual.

Levin BR, & Bull JJ. 1996. Phage Therapy Revisited: The population biology of a bacterial infection and its treatment with bacteriophage and antibiotics. The American Naturalist. 146(6): 881-898.

McNerney R. 1999. TB: The return of the phage. A review of fifty years of mycobacteriophage research. International Journal of Tuberculosis and Lung Disease. 3(3): 179-84.

McNerney R, & Traoré H. 2005. Mycobacteriophage and their application to disease control. Journal of Applied Microbiology. 99(2): 223-33.

Papke RT, & Doolittle WF. 2003. Phage evolution: New worlds of genomic diversity. Current Biology. 13(15) R606-7.

Pedulla ML, et al. 2003. Origins of highly mosaic mycobacteriophage genomics. Cell. 113(2): 171-82.

Pope WH, et al. 2015. Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. eLife. 4(e06416): 1-65.

Zelasko S, et al. 2017. Delivering phage therapy per os: Benefits and barriers. Expert Review of Anti-infective Therapy. 15(2): 167-179