Identification of novel phage

Isabel V. Abonitalla[1], Tony R. Valencia[2], and B.R. tenOever[3]

[1.] Newtown High School, Elmhurst, NY 11373. [2.] High School for Environmental Studies, NY, NY 10019. [3.] Icahn School of Medicine, NY, NY 10029

Abstract

Viruses abound in nature and show greater diversity than any other nucleic-acid based entity. To sample this diversity, we obtained common environmental samples (i.e. soil, decaying food and plants), and filtered the material to exclude all but viruses. Filtered elutes were applied to mammalian, yeast, and bacterial cultures and samples which slowed or stopped growth were further evaluated by next generation sequencing. We identified a single sample capable of slowing the division of *Staphylococcus aureus* and identified the causative agent as a novel phage most closely related to Bacillus AR9. These data illustrate viral complexity and suggest that the vast majority of viruses have yet to be discovered.

Introduction

The viruses on Earth are considered to be one of the most diverse entities. However, less than 0.01% of viruses are identified and well-characterized. These viruses can either be detrimental or beneficial to lives of human beings for they can cause widespread lethal epidemics. They can also be used as natural alternatives to antibiotics. Researchers employed the concept of the latter at a hospital in Paris when a patient had a multidrug-resistant strain of a bacteria called *Acinetobacter baumannii* [1]. The use of bacteriophages (viruses) in that case provided an organic, non-invasive solution to antibiotic resistance. Furthermore, other research, such as one completed through the observation of the bacteria *Staphylococcus* and related viruses, consistently reproduce the results and benefits of using viruses [2].

With hopes of expanding said research and more, we surveyed a wide range of samples taken from different New York City environments. We aimed to encounter a novel virus with the potential to be used as a vector (a DNA molecule used as a vehicle to artificially carry foreign genetic material into another cell) or a therapeutic (a pharmaceutical drug taken from a biological source). With each step of the research in place, we hypothesized that we would find a novel virus within our samples that would slow or stop the growth of one or more of three model systems: bacterial, yeast, or mammalian cultured cells.

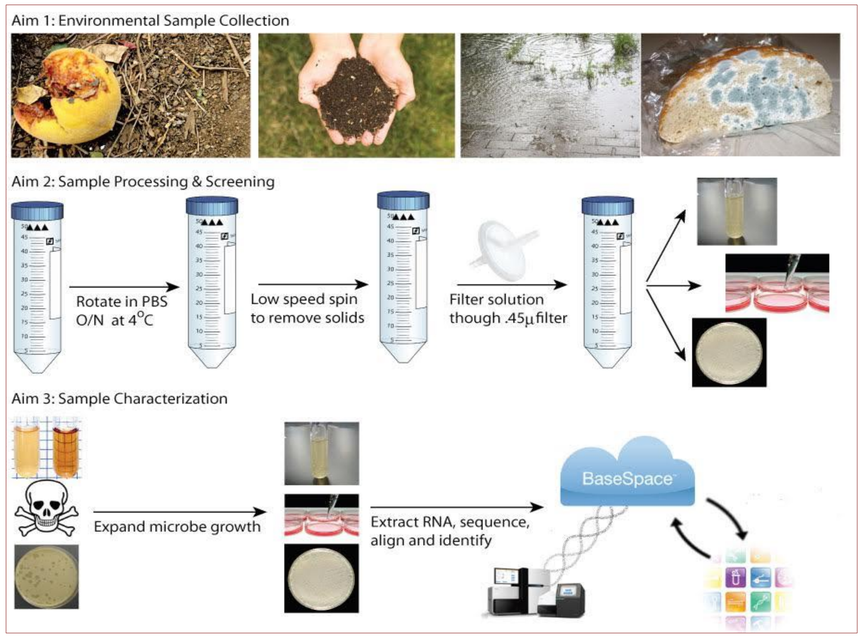
Materials & Methods

From various locations in New York City, we collected 50 possible microbe-containing samples which included the following: soil, plants, and food. After the samples were collected and sealed in tubes, we conducted differential centrifugation to filter out all but the smallest debris. Having filtrated each sample, we extracted the eluates and used a .45 micron screen that allowed only viruses to pass through. The resulting solutions were then grouped into batches of 5 and added to 10 new tubes. Having processed each sample and grouped them, we checked for evidence of viruses that may have affected one of three different model systems: a bacteria (*Staphylococcus*), a commercially used yeast (*Pichia*) and an African green monkey cell line (Vero cells). Purified phosphate-buffered saline and a soil sample spiked with replication-incompetent Adenovirus (designed to express a green fluorescent protein derived from jellyfish) served as negative and positive controls, respectively. Following the cultivation of the different hosts, we administered the samples and allowed our cultures to grow under ideal conditions for 24 hours. For the *Staphylococcus* and *Pichia*, we compared the light absorbance from the growth medium after 24 hours using a spectrophotometer machine that determined whether there was a virus that slowed down the growth of any of the cultures. As for the Vero cells, we examined each culture for cell death under a microscope.

The identification of a significant amount of prohibited growth within one of our bacteria cultures led us to the barcoding and sequencing of said sample using random cDNA library building, which is the product of the Illumina platform that allows for the expansion and amplification of genetic material. The amplified genetic material was then aligned using Bowtie-2 [3] and compared to all known viruses found on the NCBI database [4].

Results

Sequencing of our sample, A(21-24), that inhibited *S. aureus (Figure 1 and 2)* identified a 250957 nucleotide-based genome that aligned to Bacillus AR9 phage with >80% identity. In addition to polymorphisms, large genomic deletions and insertions were identified. A subset of insertions aligned to other, unrelated phages including *Yersinia* phage phiR1-37 (Figure 3).



Aim 1 *shows the collection of samples from the environment.* ***Aim 2*** *shows the processing of the samples.* ***Aim 3*** *shows the characterization of the various microbes that are grown using sequencing methods and bioinformatic systems.*

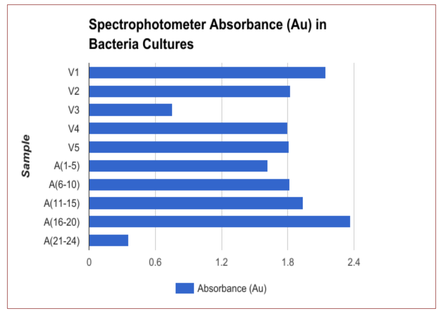


Figure 1*. The samples had varying absorbance (measured in Au units) of the light rays emitted by the Spectrophotometer. Higher levels of absorbance means that there was a significant amount of growth within the culture. The average absorbance for the bacteria cultures was 1.6434.*

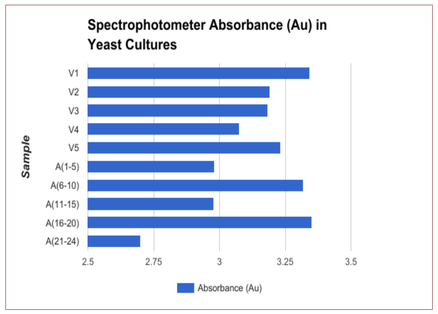


Figure 2. *The samples had varying absorbance (measured in Au units) of the light rays emitted by the Spectrophotometer. Higher levels of absorbance means that there was a significant amount of growth within the culture. The average absorbance for the yeast cultures was 3.1361.*

Acknowledgements

We thank the Pinkerton Foundation and Cold Spring Harbor Laboratory for creating and enabling this wonderful program. We also thank our mentor Benjamin tenOever for his outstanding guidance and support throughout the entirety of this project. Above all, this project could not have been completed without the aid of everyone listed above and we are beyond grateful for this experience.

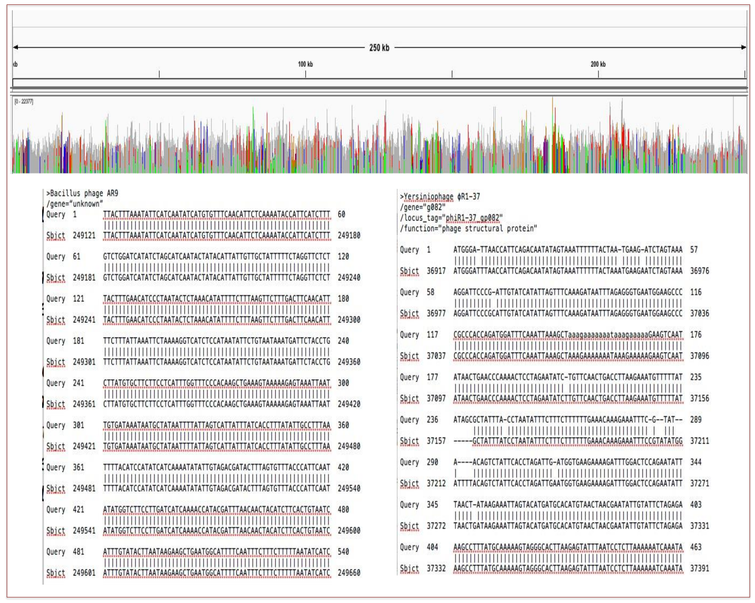


Figure 3. *The top photo shows the consensus data of our sample, whose closest evolutionary relative is Bacillus phage AR9. The bottom left photo shows the alignments of a gene shared by our phage and Bacillus phage AR9. The bottom right photo shows a unique gene that codes for a Yersinia phage structural protein and aligned with our phage. The latter gene was not found in Bacillus phage AR9.*

Discussion

The slowed division of the *S. aureus* culture with the A(21-24) sample shows that there is organic material on Earth that could be used to eliminate diseases caused by harmful bacteria, thus antibiotics may not be completely needed anymore. Furthermore, the discovery of the novel phage closely related to Bacillus AR9 helped to solidify the reality of beneficial viruses that could alleviate harmful bacteria of their tendency to cause medical issues for humans.

As for the yeast and Vero cells, the lack of identification of a virus that could stunt the growth of either cultures may be due to the types of samples and where they came from. To add to that, the Vero cells most likely were not affected by the genetic material present in our samples because our samples did not contain substances capable of harming mammalian cells -a finding of viruses capable of harming said cells would have been hazardous to the lab and researchers.

Future studies should focus on gathering larger sample sizes that come from various cities in order to have a wider diversity of possible microbe-containing samples. Also, studies should focus on having multiple types of bacteria because it will help determine if viruses can stop the growth of most known bacteria.

References

1.U.S. patient with drug-resistant bacterial infection saved by phage therapy: report. CTVNews. 2017 Apr 29 [accessed 2017 May 14]. http://www.ctvnews.ca/health/u-s-patient-with-drug-resistant-bacterial-infection-saved-by-phage-therapy-report-1.3391210

2.Kaźmierczak Z, Górski A, Dąbrowska K. Facing Antibiotic Resistance: Staphylococcus aureus Phages as a Medical Tool. Viruses. 2014;6(7):2551–2570.

3.Borozan I, Watt SN, Ferretti V. Evaluation of Alignment Algorithms for Discovery and Identification of Pathogens Using RNA-Seq. PLoS ONE. 2013;8(10).

4.Barrett T. NCBI GEO: mining millions of expression profiles--database and tools. Nucleic Acids Research. 2004;33 (Database issue).