Katherine Hand

March 12, 2021

Dr. Bradley Sekedat

UWP 101

The Secret Life of Bacteriophages: How Did They Originate?

One of life's greatest mysteries, the origin of viruses, remains unsolved. To crack the case, scientists proposed three major hypotheses. How the origins of viruses continue to remain undiscovered despite the hypotheses put forward calls for further investigation into the topic. Discovered in the early twentieth century by Frederick Twort and Felix d'Herelle (Drulis-Kawa et al., 2015), bacteriophages are considered the most successful entities (Moelling, 2012) and they help build the genomes of all species. What began as a narrow belief that viruses only prefer to genetically exchange with hosts of their superkingdom (Archaea, Bacteria, or Eukarya), studies on bacteriophages prove this untrue. The secret life of bacteriophages is seemingly more peculiar in the human gut as they coevolve with bacteria and interact with our immune system. Their distinctive interactions with white blood cells (WBC) and their hosts demonstrate a connection with their protein fold and structural make-up. This leaves us with an interesting question: where did the phages originate from? Assessing how bacteriophages can be antimorphic mutations[1] of early bacterial defense mechanisms where an accumulation of proteins that should not have bound together, bound together over time and self-infected a bacterial cell from within might be our answer.

A bacteriophage, also known as a bacteria-eater or phage, is a type of virus that infects bacteria. It has an icosahedral-shaped capsid (head) containing DNA, a neck with a collar, a tail enclosed by a sheath, and a base-plate at the end of the tail equipped with six long tail fibers that

[1] An antimorphic mutation is a dominant and negative mutation.

give the appearance of legs. Generally, bacteriophages are strain-specific. This means that it is rare to have bacteriophages that infect the majority of strains within one species and fewer phages that infect distinct species (De Sordi et al., 2019). Recognized as one of the most abundant entities in the gastrointestinal (human gut) microbiome, they are one of the least understood. The other most abundant entity is bacteria. Bacteria compete in an arms race against bacteriophages through antagonistic coevolution[2] (AC). These interactions drive microbial diversity, equilibrium, and survival (De Sordi et al., 2019).

Three major hypotheses have been seriously considered in the efforts to explain the origin of viruses. The first is the virus-first hypothesis. It alleges that viruses (or precellular, virus-like genetic elements) act as "intermediates between prebiotic chemical systems and cellular life" (Krupovic and Koonin, 2017) in the precellular world. Second is the regression hypothesis, which conveys that viruses are degenerate cells that display parasitic behavior intracellularly (or within cells) and shed many of the functions found in cellular life forms over time. Lastly, the escape hypothesis explains that viruses independently evolved from cellular genes from different domains of life that "embraced selfish replication and became infectious" (Krupovic and Koonin, 2017). However, none of these hypotheses achieved acceptance by the entire scientific world. The remainder of this paper will be dedicated to evaluating and exploring my hypothesis on the origin of bacteriophages. I will also attempt to assess how the functions and behaviors of bacteriophages and WBC are similar. This attempt will allow for the formation of an explanation as to why the origins of bacteriophages might demonstrate how they infect other domains of life aside from their hosts.

[2] Antagonistic coevolution between bacteria and bacteriophages is the reciprocal evolution of bacterial resistance and phage infectivity (Scanlan 3).

Acknowledging the observation of bacteriophage adaptation to bacterial resistance in AC uncovers that bacteriophages with the most mutations have the ability to infect past, present, and future bacterial clones (De Sordi et al., 2019). However, it is unclear which genomic mutations are associated with bacteriophage infectivity. In an attempt to narrow down what mutations are responsible, Sordi, Lourenço, and Debarbieux performed a time-shift interaction study to determine which of the mutations amassed before or after bacterial resistance. They isolated three types of bacteriophage clones: one unable to infect a present bacterial clone, and two that have adapted to infect past and present bacterial clones. By day 21 of the isolated study, all bacteriophage clones had the ability to infect past, present, and future bacterial clones. They argue that the arms race between bacteria and bacteriophages was initially characterized by a "rapid development of bacterial resistance" (De Sordi et al., 2019) and bacteriophage adaptation came later.

The result of this study suggests that bacteriophages initially attacked bacteria. In response, bacteria had to develop resistance. Then, bacteriophages needed to mutate and adapt to infect bacteria again. These findings coincide with my hypothesis since the past bacterial clones were unable to resist the phage implying it did not anticipate the attack. The phages were somehow equipped with the proper tools to infect bacteria. If a phage is an accumulation of proteins bound within a bacterial cell, it could be possible that the phage was born intracellularly with the right proteins to stage the attack. Especially if these proteins are bound over time by an irreversible and negative mutation that has been copied and passed on to other cells.

Further evidence of bacteria-bacteriophage coevolution shows growing evidence towards bacteriophages regularly interacting with immune cells of the innate immune system[3] by

[3] The innate immune system protects the Eukaryotic host from invading microorganisms and bacteriophages (Carroll-Portillo and Lin 1).

Carroll-Portillo and Lin in their 2019 study. Through the activation of the immune system, Carroll-Portillo and Lin suggest that bacteriophages play a major role in the function of the eukaryotic host. Bacteriophages travel outside the mucosal environments and into the bloodstream, spleen, kidney, liver, and brain through mechanisms of phage transcytosis[4] such as free uptake[5], the trojan horse[6], or by crossing via a leaky gut[7] to reach immune cells (Carroll-Portillo and Lin, 2019). As a result, several questions are posed: why are bacteriophages interacting with the metazoan immune system if it is not their host? Could bacteriophages be acting reminiscent of animal viruses? To tackle these questions, I will assume a biochemical approach to the problem. First, I will observe bacteriophage behavior. Then, break it down to its structures and components before putting the virus back together. I will primarily focus on virus-host interactions, the viral capsid (head), and phage tail (leg) fibers.

While one could assume that bacteriophages would only interact and infect its host, they would be incorrect. Recent studies blur the traditional understanding of virus-host interactions. A comparative genomic approach was used to explore proteomes[8] grouped into fold superfamilies (FSFs). Proteomes were analyzed with cross-superkingdom genetic exchanges occurring between archaea, bacteria, and eukaryotes with the three virus groups: archaeoviruses, bacterioviruses, and eukaryoviruses. Protein domain groups contain interlockings of amino acid side-chains found in their cores. These serve as “fingerprints" to distinguish the fields among domains of any superfamily. Tracing the spread of each viral FSF in cellular proteomes helps to

---

[4] Transcytosis is a form of transportation to carry particular molecules across a cell.
[5] Free uptake is a mechanism where the phage alone is endocytosed (process of how substances are brought into the cell) and transported across the gastrointestinal epithelium (tissue that forms the outer layer of a surface) (Carroll-Portillo and Lin 3).
[6] Trojan horse is a mechanism where a phage residing within a bacterium transports across a barrier (Carroll-Portillo and Lin 5).
[7] Crossing via a leaky gut is a mechanism where inflammation or injury results in an impaired barrier function which allows for passive transit of phage through the epithelium (Carroll-Portillo and Lin 3).
[8] Proteomes are proteins expressed by an organism.

yield insight into the genetic exchanges between viruses and cells. It demonstrated a cross-superkingdom genetic exchange between bacterioviruses and eukaryotes and eukaryoviruses and bacteria (Malik et al., 2017).

To determine whether the exchanges occurred either from virus-to-cell or cell-to-virus Horizontal Gene Transfer[9] (HGT) events, two factors were considered. The first was through the biochemical function of an FSF in question. For example, "if an FSF involved in capsid assembly (a virus hallmark function) was detected in only a few cellular proteomes (e.g. <1%)" (Malik et al., 2017), then it was determined that this exchange occurred by a virus-to-cell event. The other factor demonstrated a cell-to-virus HGT event if the FSF was widespread in cellular proteomes (Malik et al., 2017). There is also an exception to these factors: that being the presence of cell-like FSFs in the proteomes of all three virus groups. This suggests that there was a "cellular coexistence between viral and cellular ancestors prior to diversification of modern life" (Malik et al., 2017). However, how is it possible that viruses are making cross-superkingdom exchanges? What does this say about their origins? How does this reveal that bacteriophages are acting reminiscent of animal viruses? To clarify these questions, however, we must first focus on specific structures of the phage.

The unique icosahedral-shaped capsids of the bacteriophage make them easily recognizable among other viruses. In fact, one of the most signature features of a virus is its capsid, and the origins of the viruses are inseparable from the emergence of this feature. Curious about the evolution of major virion proteins, Krupovic and Koonin completed a comprehensive sequence and structure analysis. They found that the major virion proteins evolved on about twenty independent occasions. Although viral replication proteins have no closely related

[9] The movement of genetic material from parent to offspring.

homologs[10] to cellular organisms, numerous structural comparisons uncovered unexpected similarities between viral capsid proteins (CPs) and cellular proteins from viruses infecting hosts of different domains of life (Krupovic and Koonin, 2017).

Recognizing these similarities led Krupovic and Koonin to suggest that viruses evolved on multiple, independent occasions through the recruitment of diverse host proteins that eventually became major virion components. This scenario works in support and the extension of my hypothesis about the origins of viruses. As bacteria reproduced and performed gene transfers with other bacterial cells, the mutated assembly of proteins (that would eventually evolve into the emergence of a virus) was either shared with other cells or shared proteins from the other cells bound to it. To break down the capsid further, it is essential to investigate specific proteins within the viral head and compare it to their hosts.

One protein of interest, Hoc[11], belongs to the Immunoglobulin Superfamily[12] (IgSF). This protein was discovered on the surface of the mature bacteriophage T4 head. Hoc has a dumbbell-like shape protruding about 6nm away from the capsid surface (Putra and Lyrawati, 2020). However, it is not present in other phages, for instance, bacteriophage T2. Scientists Bateman, Eddy, and Mesyanzhinov express that Hoc is non-essential to the structure or viability for T4, and its exact function is unknown. In a study performed by Putra and Lyrawati, they reveal that Hoc "probably evolved as a form of adaptation of T4 bacteriophages to avoid immune system recognition, thereby allowing them to survive inside eukaryotes" (Putra and Lyrawati, 2020). Nonetheless, Hoc's exact function in the phage and how it got there remains unknown.

---

[10] Homologs are homologous chromosomes, or 2 alleles of chromosomes that result into two copies of each allele after DNA replication.
[11] Hoc is a highly immunogenic outer capsid protein (Bateman et al).
[12] Members of IgSF are important in vertebrate immune systems. They include: antibodies, T-cell receptors, major histocompatibility antigens, and cell adhesion molecules. It is unknown how the IgSF domain arose (Bateman et al).

Bateman et al. also observed that the proteins in bacteriophage T4 are most similar in sequence and enzymatic mechanisms of eukaryotic rather than prokaryotic[13] homologs (Bateman et al., 1997). They raise the possibility: could T4 be a late evolutionary invention constructed in part from eukaryotic genes?

Before answering this question, it is essential to discuss phage tail (leg) fibers. When a phage encounters a cell, it delivers its genetic material into the host cytoplasm. On bacteriophage T7, the six tail fibers are folded against the capsid and extend when in contact with a host cell surface (Hu et al., 2013). Hu et al. used a cryo-electron tomography[14] (cryo-ET) to capture bacteriophage T7's virions at different, successive stages of infection on E. coli. As tail fibers bind to the cell, the phage "walks" across the cell's surface in search of a receptor for its tail. Once located, all the fibers rotate downward, and the infection begins (Hu et al., 2013).

So what does phage "walking" and the existence of the Hoc protein have to do with Eukarya and the human gut? When examining the six tail fibers of the bacteriophage, what other cellular organism shares a similar ability to "walk?" It was initially believed that lymphocytes move over endothelial cells as flattened bodies (Alon et al., 2009). However, Alon et al. observed that they use ventral adhesive filopodia[15], or legs, to rapidly crawl and scan endothelial cells via millipede-like locomotion. Significantly, this shows that both the phage and the lymphocyte are similar by their ability to walk and scan for a location. It is possible that lymphocytes gained this ability from the phages through a virus-to-cell event.

Still, why are phages making cross-superkingdom exchanges, and how are they able to do this? It is probable that a negative mutation occurred within a bacterial cell. Proteins that should

---

[13] Prokaryotes are single-celled organisms that lack a nucleus.
[14] Cryo-electron tomography provides a three-dimensional (3D) structure of infected cell complexes in near-native, frozen-hydrated states at ~4-nm resolution (Hu et al).
[15] Adhesive filopodia is defined as a protrusion extending from the lymphocyte body in direct contact with the substrate.

not have bound together, bound together. As bacteria engaged in HGT and reproduction, more bacteria contained this hidden negative mutation. As time passed, more proteins would bind. As for making cross-superkingdom exchanges, it is known that eukaryotic cells engage in phagocytosis to engulf large particles. This is how eukaryotic cells acquired bacterial endosymbionts, for instance, the mitochondria (Yutin et al., 2009). Therefore, what is not to say that a eukaryotic cell engulfed a bacteria carrying the protein mutation? There is the chance that this is how bacteriophage T4 gained the Hoc protein and other similar eukaryotic sequences. This could serve as an explanation for how the phages are performing exchanges with Eukarya. Because of the similar eukaryotic sequences, bacteriophages may sense a connection or recognize eukaryotic cells. This could reveal why they interact with innate immune cells rather than acting reminiscent of animal cells. As for T4 being a late evolutionary invention constructed in part from eukaryotic cells, that may provide a solid answer to why bacteriophage T2 lacks Hoc. T2 was likely not provided the opportunity to bind a Hoc protein to its capsid because it was not engulfed by a eukaryotic cell.

It is the lack of research and knowledge on bacteriophages that makes these entities so mysterious to us. Because of the compelling similarities shared between phages and immune cells, it might be pure irony or perhaps something more. Therefore, it is crucial to recognize that these phages may hold the key to unlocking the answer to viral origins. Future research on bacteriophages may provide solid grounds to further investigate my hypothesis. Bacteriophages may have begun as an antimorphic mutation within bacterial cells leading to the event that certain proteins bound together that should not have bound. Over time, the mutated proteins were shared with other bacterial cells through HGT and as it evolved, mutated, or was phagocytosed, more proteins bound. The idea of a bacterial cell containing the mutated proteins being

phagocytosed by a eukaryotic cell could explain how bacteriophage T4 shares so many similarities with Eukarya. Moreover, if immune cells gained the ability to walk, then it might also be possible that phages gave rise to that ability by interacting with WBC. Thus, the secret life of bacteriophages is highly complex, beginning from within bacteria.

Works Cited

Alon, Ronen, et al. “Lymphocyte Crawling and Transendothelial Migration Require Chemokine Triggering of High-Affinity LFA-1 Integrin.” *Immunity*, 5 Mar. 2009, www.cell.com/immunity/fulltext/S1074761309001022.

Bateman, Alex, et al. “A Member of the Immunoglobulin Superfamily in Bacteriophage T4.” *Virus Genes*, Springer US, 1 Jan. 1997, rd.springer.com/article/10.1023/a:1007977503658.

Carroll-Portillo, Amanda, and Henry C Lin. “Bacteriophage and the Innate Immune System: Access and Signaling.” *Microorganisms*, MDPI, 28 Nov. 2019, www.ncbi.nlm.nih.gov/pmc/articles/PMC6956183/.

De Sordi, Luisa, et al. “‘I Will Survive’: A Tale of Bacteriophage-Bacteria Coevolution in the Gut.” *Gut Microbiomes*, 10:1, 92-99, 2019, www.tandfonline.com/doi/full/10.1080/19490976.2018.1474322.

Drulis-Kawa, Zuzanna, et al. “Bacteriophages and Phage-Derived Proteins--Application Approaches.” *Current Medicinal Chemistry*, Bentham Science Publishers, 2015, www.ncbi.nlm.nih.gov/pmc/articles/PMC4468916/.

Hu, Bo, et al. “The Bacteriophage t7 Virion Undergoes Extensive Structural Remodeling during Infection.” *Science (New York, N.Y.)*, U.S. National Library of Medicine, 1 Feb. 2013, www.ncbi.nlm.nih.gov/pmc/articles/PMC3873743/.

Krupovic, Mart, and Eugene V. Koonin. “Multiple Origins of Viral Capsid Proteins from Cellular Ancestors.” *PNAS*, National Academy of Sciences, 21 Mar. 2017, www.pnas.org/content/114/12/E2401#ref-85.

Malik, Shahana S., et al. “Do Viruses Exchange Genes across Superkingdoms of Life?”

Frontiers, Frontiers, 16 Oct. 2017, www.frontiersin.org/articles/10.3389/fmicb.2017.02110/full.

Moelling, Karin. "Are Viruses Our Oldest Ancestors?" EMBO Reports, Nature Publishing Group, Dec. 2012, www.ncbi.nlm.nih.gov/pmc/articles/PMC3512416/.

Putra, Ramendra Dirgantara, and Diana Lyrawati. "Interactions between Bacteriophages and Eukaryotic Cells." Scientifica, Hindawi, 9 June 2020, www.ncbi.nlm.nih.gov/pmc/articles/PMC7301238/.

Scanlan, Pauline D. "Bacteria–Bacteriophage Coevolution in the Human Gut: Implications for Microbial Diversity and Functionality." *Trends in Microbiology*, Elsevier Current Trends, 22 Mar. 2017, www.sciencedirect.com/science/article/pii/S0966842X17300495?casa_token=cysIDnJgTvsAAAAA%3Az2BwBy8jv2LMK_898yGvIhoBQQNJhTswHNI2VVlpbYNv3cDKPr-pVymO_yRARp1KsFxy-FiXOw.

Yutin, Natalya, et al. "The Origins of Phagocytosis and Eukaryogenesis." *Biology Direct*, BioMed Central, 26 Feb. 2009, www.ncbi.nlm.nih.gov/pmc/articles/PMC2651865/.