
THE BOOLEAN
SNAPSHOTS OF DOCTORAL RESEARCH AT UCC

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

The Cosmopolitan Gut Virus crAssphage

Linda Smith[1,*]

[1] *Department of Microbiology, University College Cork, Cork, Ireland*
[*] *Corresponding author: linda.smith@ucc.ie*

Abstract

CrAss-like phages are a diverse group of mostly uncultured bacterial viruses that are highly abundant in the mammalian gut and other habitats. First identified in metagenomic sequences from human faeces in 2014, crAss-like phages were predicted to infect members of the phylum Bacteroidetes. Later work resulted in the isolation of the first cultured representatives, the confirmation of a Podoviridae-like morphology and a proposal to classify uncultured crAss-like phages under a novel taxonomic group. The International Committee on Taxonomy of Viruses (ICTV) has recently acknowledged the creations of a new order Crassvirales, comprising at the moment four new families, ten new subfamilies, 42 new genera and a total of 73 new species. Many unanswered questions remain about this ubiquitous yet enigmatic gut virus which I aim to approach in my research through computational analysis of metagenomic sequences.

Keywords: bioinformatics, microbiology.

> *We are all of us walking communities of bacteria. The world shimmers, a pointillist landscape made of tiny living beings.*
>
> — Lynn Margulis

Our modern microbiome

Most people today have heard of the microbiome - a collection of microscopic organisms composed of bacteria, viruses and fungi that live inside and on us. Colonizing every nook and cranny of our bodies, microbes enjoy inhabiting our skin, teeth, airways and especially the inner lining of our gastrointestinal tract. Scientists estimate that the ratio of human cells to microbes in our bodies is one-to-one. Despite being a relatively young field, microbiome research is revealing the multitude of ways microbes interface with us. Bidirectional communication is happening between our bodies and our bugs. They impact our digestion and nutrition, modulate our immune systems, our appetite control and energy balance. They influence our cognition, behaviour, and even our happiness. A dysregulated microbiome has been implicated in disorders such as inflammatory bowel disease, obesity, major depressive disorder, Alzheimer's and

diabetes. Faecal microbiome transplants (FMTs) hold promise as potential therapeutics against many diseases. Experiments have shown that transplants from young to old mice can reverse the hallmarks of aging, illuminating the significance the microbiome plays in the vitality and health of the overall system. Companies target us with products to "feed our beneficial bacteria". It's true, we can greatly improve our health by making peace with our gut residents. We can directly influence our gut bacteria dwellers through our diet, but you've probably never considered the extra dimension of the viruses that live inside our bacteria. Many human diseases are associated with not only changes in the bacteriome composition of our gut but also that of the virome.

The war raging in our gut

No the war isn't between your bowel and those garlic cheese chips, it's happening between bacteria and the viruses that inhabit them. Bacteriophages, meaning bacteria eaters, are the notorious viruses that prey on bacteria. Harmless to human cells, these distinctive icosahedral (20 faced) apex predators are highly adapted to hijacking and killing bacteria. Bacteria-virus ecology is a rather niche topic but an important one considering phages impact the numbers and evolution of bacteria in our gut. The majority of viruses in a healthy gut (∼90%) are bacteriophages and are collectively known as the human gut phageome. I am lucky to be part of the gut phageomics group in UCC who are at the forefront of researching the dynamic interactions of phages in the human gut.

Bacteriophage life cycle

Some phages live a double-life, as shown in Figure 2. They have two states of existence: a dormant, or lysogenic state where they can incorporate their DNA into the host bacteria's DNA, piggybacking off the bacteria's resources and reproductive life cycle until they choose to switch into an active or lytic state. In their lytic state, the phage takes over the host bacterial cell, transforming it into a phage assembly factory, assembling an army of phage structures to burst from the host bacteria, almost always killing it, to invade new bacteria and begin the cycle anew. Lytic phages are of focus for use in phage therapy - a targeted kill missile against certain pathogenic species of bacteria. Phages are of special interest to microbial ecologists since they drive the evolution of bacteria through their efficient packaging and delivery of genes in a process called transduction. When existing as a part of the host bacteria's DNA, lysogenic phages get replicated with each successive cell division cycle of the bacteria. This segment of incorporated phage DNA in a bacterial chromosome is called a prophage. Upon sensing conditions either inside the bacterial cell or outside it, the prophage DNA can become expressed or induced, switching gears into the lytic cycle. One study showed that different foods and chemicals we ingest can either induce the prophages into their active lytic cycle or suppress them in their lysogenic state.

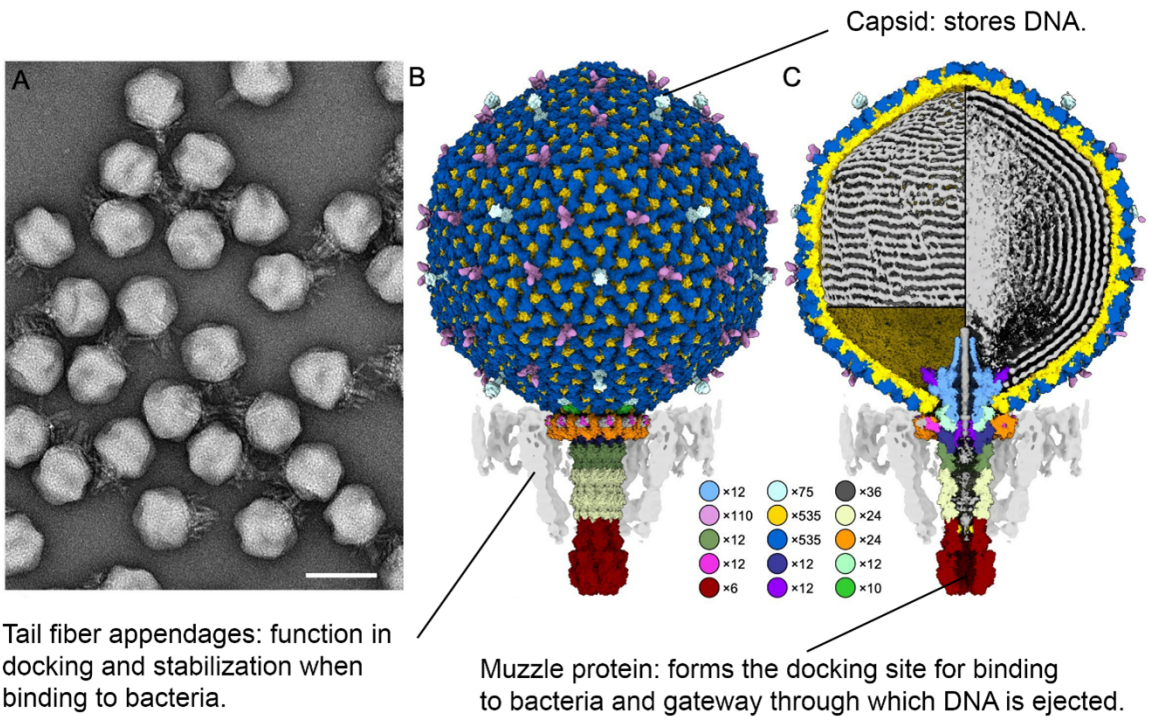

Figure 1: Transmission electron microscope (TEM) image of ϕcrAss001 virions (A), cryo-EM structure (B,C). Cryo-EM morphology resolved by York Structural Biology Lab, Bayfield, Shkoporov et al.[8]

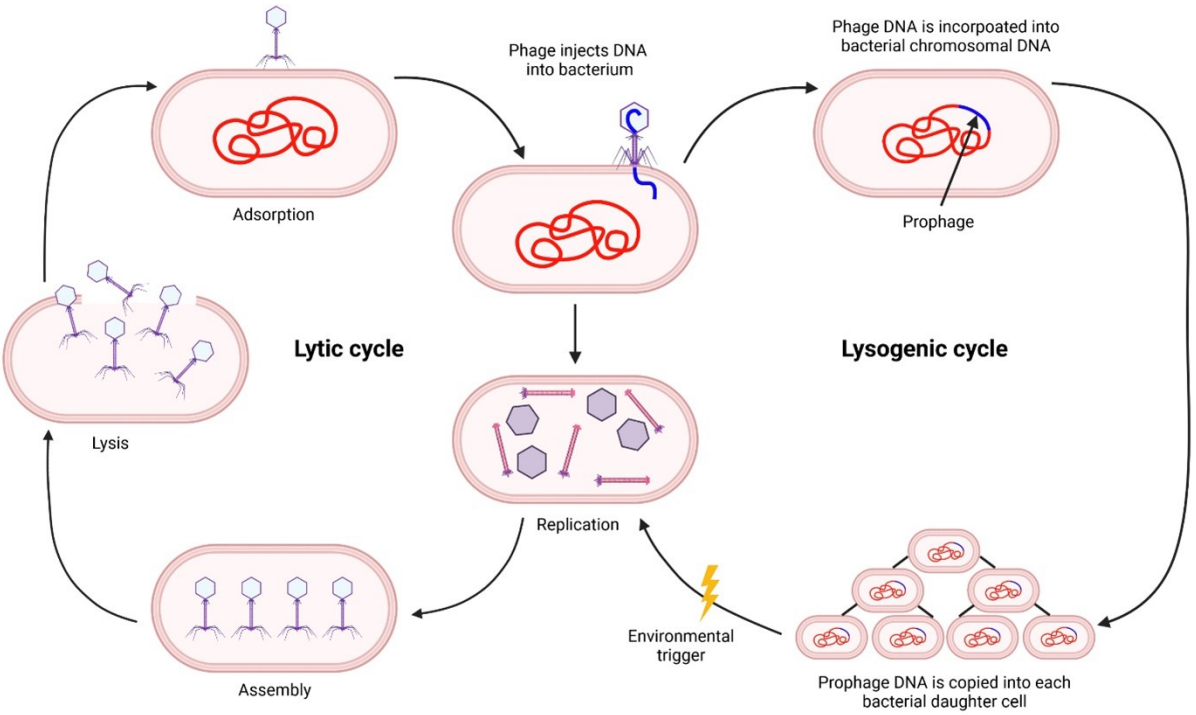

Figure 2: Bacteriophage life cycle

Phages are gene traffickers

A fascinating aspect to phage biology is their gene trafficking capabilities. Phages can change the physiology of their host bacteria by introducing new genes into the bacteria's genome. In this way, phages can shape bacterial communities, conferring genetic toolkits to their bacteria which make them more adaptable to their environment. This is known as horizontal gene transfer and occurs by transduction. One recently discovered form called lateral transduction, is calling into question concepts in microbiology previously thought to be established. Previously it was thought that "mobile genetic elements" existed within the bacteria's DNA and they could "jump" or move between bacteria. Evidence is unearthing phages to be the gene movement masterminds all along.

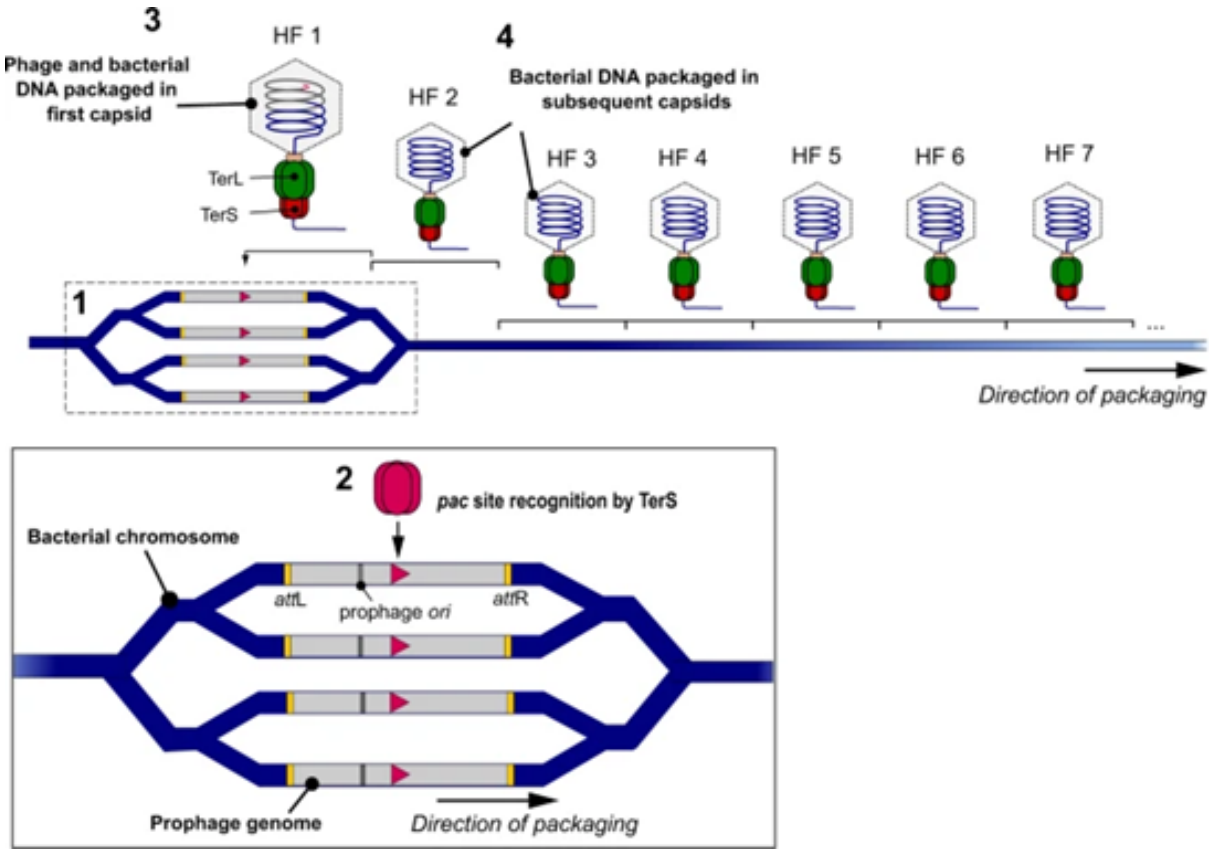

Figure 3: Lateral transduction of bacterial chromosomal DNA. Humphrey et. al

In lateral transduction, phages are able to package up bacterial DNA in-situ into their capsids and shuttle them out of the cell. It looks like a factory assembly line! A bacterial gene located next to the prophage integration site can theoretically be delivered "around the world" to the wider bacterial community. By this mechanism, phages can share antibiotic resistance genes and pathogenic genes to bacteria. A direct example of this is when the CTXϕ bacteriophage introduces the cholera toxin gene into the genome the bacteria *Vibrio cholera*, causing it to become especially virulent.

Ancient associations

As far back as we can trace life on Earth, to the last universal common ancestor (LUCA), viruses have existed. Bacteria and phages have co-evolved for billions of years and studying their interactions have yielded us jackpot advances in biotechnology such as CRISPR - a bacterial defence mechanism against phages. They are also the most abundant and diverse biological entities found on Earth. Despite being such a simple entity - consisting of just a head (capsid) storing their genetic material and a tail to which they bind to bacteria - scientists continue to be surprised by the novel genetic mechanisms and complexities by which phage operate. Phages serve an exciting glimpse into the diverse and ancient microbial world time and time again.

The benign bacteriophage

CrAssphage, the most abundant virus in the human gut, was discovered by chance in 2014. Named after the cross-assembly algorithm used to discover them, not their location of origin, the discovery of crAssphage is a triumph of modern DNA sequencing technology. Before the dawn of high-throughput sequencing and massive computing capacity, scientists would have to isolate the virus in the lab. To do this for a phage you would first need to know its bacterial host, try to grow it in the lab and then study its phage. A daunting and time consuming task. Now we can take a sample from any environment, be it seawater, soil, sewage or a stool sample, and sequence every DNA molecule present in that sample. Like a puzzle, we can assemble the pieces of the DNA to get a snapshot of the community of microorganisms present there. This field is known as metagenomics and is currently revolutionizing microbiology. After pooling the gut metagenomes of 4 pairs of female twins and their mothers, and piecing their microbial genetic puzzles together, Dutilh et al[3] found a 97,000 base pair segment of DNA that was in very high abundance and was shared across all the subjects. After closer examination of this DNA segment they discovered it was a bacteriophage. This is the story of how the prototypical crAssphage was found in the human gut. Since its discovery, the famous crAssphage has been detected in human faecal samples all around the world, in populations as remote as rural Malawi and from the Amazonas of Venezuela. 98-100% of healthy adults from Western samples carry at least 1 type of crAssphage. It is found in all age groups (from as young as 1 year old to 65+) and it is thought to be acquired during early childhood. CrAssphage is also a benign gut inhabitant as no associations between it and diet or health have been found. Currently in industry it is being used as a signature of human faecal contamination in wastewater sources and other environments. First isolated in UCC's gut phageomics lab, ϕcrAss001 was found to infect a bacteria in the phylum of Bacteroides, one of the most abundant bacterial members of our gut, who carry out important functions such as metabolizing carbohydrates and providing nutrition and vitamins to us and to the other gut residents.

Viral dark matter

Advances in sequencing technology have led to an enormous amount of data being generated from sampled environments. One of the first approaches you would use to better understand your sequencing data would be to align your sequences against a database of known reference genomes. Often a segment of your sequence will have a similarity or a hit against a genome or gene in a database. From this you can infer the species your sequence came from or at the very least if it's bacterial or viral. The term "viral dark matter" has emerged to describe the viral sequences for which we have no known reference in our databases. Our databases are rapidly expanding with hundreds of thousands of viral sequences which we have never encountered before, isolated in the lab or have any inkling of their role in the earth's ecosystem. The potential for the discovery and classification of new viruses is huge, with metagenomic sequence analysis becoming the primary mode of virus discovery. The current pace of viral discovery is exceeding our ability to classify these viruses and put them in their taxonomic ranks.

A place for crAssphage in global viral taxonomy

This is where my PhD comes in. Using my laptop, I am going to search through diverse and obscure databases online from multiple sources – human and mammalian gut metagenomes, soil, sewage, and seawater samples - for viral sequences similar to the crAss-like viruses that have been previously found in the human gut. In the process I hope to discover new crAssphage species which we can recruit into a revised taxonomic grouping of all known global crAssphage sequences. A crAssphage was recently found in the faecal sample of a cat! Many unanswered questions remain about crAssphages. Do they exist in all mammalian gut metagenomes or just animals that have direct contact with humans? Why are they so abundant in humans? What type of life cycles do they have? Do crAssphages infect bacteria outside of the *Bacteroides* genus? As well as creating a revised taxonomy of crAssphages and gaining a big picture view of their place in the global viral taxonomy, we hope that by accumulating and sharing a large database of crAss-like sequences these unanswered questions will be approachable for anyone who wishes to probe the data. Finding how the most abundant human virus - crAss-like phages - relate to existing viruses and interact with their host bacteria will have significance for virologists and the wider microbial community.

Acknowledgements

Thanks to my supervisors Andrey Shkoporov and Colin Hill and to Science Foundation Ireland for funding my research.

Declaration of interests

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. All ethical guidelines relating to the research and publication

process were adhered to throughout this study.

References

[1] Ilseung Cho. Martin j. *Blaser-The Human Microbiome: at the interface of health and disease-Nat Rev Genet*, 13(4):260–270.

[2] André M Comeau, Graham F Hatfull, Henry M Krisch, Debbie Lindell, Nicholas H Mann, and David Prangishvili. Exploring the prokaryotic virosphere. *Research in microbiology*, 159(5):306–313, 2008.

[3] Bas E Dutilh, Noriko Cassman, Katelyn McNair, Savannah E Sanchez, Genivaldo GZ Silva, Lance Boling, Jeremy J Barr, Daan R Speth, Victor Seguritan, Ramy K Aziz, et al. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nature communications*, 5(1):1–11, 2014.

[4] Bas E Dutilh, Arvind Varsani, Yigang Tong, Peter Simmonds, Sead Sabanadzovic, Luisa Rubino, Simon Roux, Alejandro Reyes Muñoz, Cédric Lood, Elliot J Lefkowitz, et al. Perspective on taxonomic classification of uncultivated viruses. *Current opinion in virology*, 51:207–215, 2021.

[5] Robert A Edwards, Alejandro A Vega, Holly M Norman, Maria Ohaeri, Kyle Levi, Elizabeth A Dinsdale, Ondrej Cinek, Ramy K Aziz, Katelyn McNair, Jeremy J Barr, et al. Global phylogeography and ancient evolution of the widespread human gut virus crassphage. *Nature microbiology*, 4(10):1727–1736, 2019.

[6] Emma Guerin and Colin Hill. Shining light on human gut bacteriophages. *Frontiers in Cellular and Infection Microbiology*, 10:481, 2020.

[7] Mart Krupovic, Valerian V Dolja, and Eugene V Koonin. The luca and its complex virome. *Nature Reviews Microbiology*, 18(11):661–670, 2020.

[8] Andrey N Shkoporov, Ekaterina V Khokhlova, C Brian Fitzgerald, Stephen R Stockdale, Lorraine A Draper, R Paul Ross, and Colin Hill. ΦCrAss001 represents the most abundant bacteriophage family in the human gut and infects *Bacteroides* intestinalis. *Nature communications*, 9(1):1–8, 2018.

[9] Michele Zuppi, Heather L Hendrickson, Justin M O'Sullivan, and Tommi Vatanen. Phages in the gut ecosystem. *Frontiers in Cellular and Infection Microbiology*, page 1348, 2022.