CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Gut bacteriophage: Current understanding and challenges
Authors	Sutton, Thomas D. S.;Hill, Colin
Publication date	2019-11-29
Original Citation	Sutton, T. D. S. and Hill, C. (2019) 'Gut Bacteriophage: Current Understanding and Challenges', Frontiers in Endocrinology, 10, 784 (18pp). doi: 10.3389/fendo.2019.00784
Type of publication	Article (peer-reviewed)
Link to publisher's version	10.3389/fendo.2019.00784
Rights	© 2019, Sutton and Hill. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms. - http://creativecommons.org/licenses/by/4.0/
Download date	2026-11-16 11:00:16
Item downloaded from	https://hdl.handle.net/10468/9381

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

frontiers in Endocrinology

REVIEW
published: 29 November 2019
doi: 10.3389/fendo.2019.00784

Gut Bacteriophage: Current Understanding and Challenges

*Thomas D. S. Sutton and Colin Hill**

APC Microbiome Ireland and School of Microbiology, University College Cork, Cork, Ireland

The gut microbiome is widely accepted to have a significant impact on human health yet, despite years of research on this complex ecosystem, the contributions of different forces driving microbial population structure remain to be fully elucidated. The viral component of the human gut microbiome is dominated by bacteriophage, which are known to play crucial roles in shaping microbial composition, driving bacterial diversity, and facilitating horizontal gene transfer. Bacteriophage are also one of the most poorly understood components of the human gut microbiome, with the vast majority of viral sequences sharing little to no homology to reference databases. If we are to understand the dynamics of bacteriophage populations, their interaction with the human microbiome and ultimately their influence on human health, we will depend heavily on sequence based approaches and *in silico* tools. This is complicated by the fact that, as with any research field in its infancy, methods of analyses vary and this can impede our ability to compare the outputs of different studies. Here, we discuss the major findings to date regarding the human virome and reflect on our current understanding of how gut bacteriophage shape the microbiome. We consider whether or not the virome field is built on unstable foundations and if so, how can we provide a solid basis for future experimentation. The virome is a challenging yet crucial piece of the human microbiome puzzle. In order to develop our understanding, we will discuss the need to underpin future studies with robust research methods and suggest some solutions to existing challenges.

Keywords: virome, bacteriophage, microbiome, phage-host interactions, microbiota

OPEN ACCESS

Edited by:
Pierre De Meyts, Catholic University of Louvain, Belgium

Reviewed by:
Andrzej Gorski, Ludwik Hirszfeld Institute of Immunology and Experimental Therapy (PAN), Poland
Jean Albert Boutin, Servier (France), France

****Correspondence:***
Colin Hill
c.hill@ucc.ie

Specialty section:
This article was submitted to Molecular and Structural Endocrinology, a section of the journal Frontiers in Endocrinology

Received: *09 August 2019*
Accepted: *28 October 2019*
Published: *29 November 2019*

Citation:
Sutton TDS and Hill C (2019) Gut Bacteriophage: Current Understanding and Challenges. Front. Endocrinol. 10:784. doi: 10.3389/fendo.2019.00784

INTRODUCTION

The human gastrointestinal tract (GIT) is a complex environment containing billions of microorganisms (1). Changes in oxygen concentration, pH, nutrient availability, water availability and bile salts shape the relative abundance of microorganisms from all domains of life (fungi, protists, bacteria, and archaea) (2–5). Of these microorganisms, bacteria are by far the most characterized, making up the vast majority of the DNA sequences and biomass (6, 7). This bacterial community also plays a central role in normal physiology of the mammalian gut by facilitating metabolic functions, protecting against pathogens and modulating the immune system (8–10). Similarly, alterations in the composition and abundance of this bacterial community are closely associated with diseases such as irritable bowel syndrome (IBS), inflammatory bowel disease (IBD), colorectal cancer (CRC), *Clostridium difficile* infection (CDI), obesity, and neurological disorders (11–15). However, the forces that shape the composition of these bacterial communities remain poorly understood and this has slowed the development of microbiome based therapeutics and biomarkers.

Bacteriophage (phage) are viruses that infect prokaryotic hosts and play crucial roles in shaping the composition and diversity of bacterial communities in many environments, facilitating horizontal gene transfer, and nutrient turnover through continuous cycles of predation and

coevolution (16–18). To date, the majority of viral metagenome (virome) research has been focused on environmental communities such as those in the ocean (19, 20). In this environment, the virome is central to the movement of dissolved organic matter across trophic levels of the ocean food chain and between the surface and the depths of the water column (16, 21). A growing body of evidence also suggests the virome can shape the functional capacity of host communities encoding functions such as photosynthetic genes in the photic zones of the ocean (22) and bacterial virulence factors in pathogenic bacteria (23).

Phage make up the vast majority of the viral component of the gut microbiome (24). They are also believed to play a key role in shaping the composition and function of the human gut microbiome in both health and disease (25–27). However, despite being highly abundant in the gut ($>10^{10}$ g^{-1}) (28, 29) and having considerable impacts on microbial ecosystems, they remain one of the least understood members of the gut microbiome. Early sequencing studies of the human gut virome estimated that it was dominated by novel sequences, with only 41% sharing homology to databases (30). However, as sequencing platforms and library preparation methods improved and yielded a more detailed view of the virome, this unknown majority or "viral dark matter" was found to make up an even greater proportion of the virome, lowering the identifiable fraction to as little as (1–14%) (31).

Since phage were first identified by Frederick Twort in 1915 (32), culture-based methods such as plaque assays have been used to screen and quantify phage titers from many environments. Today, these methods still play a central role in identifying phage which target specific bacteria and have contributed to our understanding the mechanics of phage host interactions and replication cycles. However, as the vast majority of phage-host pairs in the gut are unknown, these methods are not suited to large-scale characterization of a complex ecosystem such as the human gut. Additionally, many of bacteria in the human gut are not routinely cultivated, despite recent advances (33). As a result, virome studies lean heavily on sequencing based metagenomic approaches to investigate gut phage communities and to try to understand their role in shaping the gut microbiome (31). This involves sequencing the total viral DNA and RNA from a community following physical separation from the bacterial component, using assembly software to recreate the viral genomes within that community and characterizing abundance and function of those genomes. However, many sequence-based virome studies exclude viral dark matter from analysis, working largely with a small fraction of known phage sequences (usually 1–14% of the dataset). This can have profound implications for the conclusions drawn from these studies, as changes in the known fraction may not reflect changes in the virome as a whole. As a result, database-independent analysis methods are increasingly being used which include both known and unknown fractions of the virome (29). However, high levels of inter-individuality make biological signals across virome studies difficult to detect (24, 34, 35). Furthermore, virome studies are particularly susceptible to methodological bias due to difficulties in benchmarking *de novo* bioinformatic tools and the dominance of unknown sequences in virome datasets (36–38).

We will discuss phage of the human gut virome and their interactions with the microbiome. We will also highlight how little we know about their role in human health. Finally, we will discuss critical areas of virome analysis methods which must be addressed and improved upon if we are to fully understand the role of phage in shaping the microbiome and human health.

HOW PHAGE INTERACT WITH BACTERIAL HOSTS

Phage Infection Cycles

As obligate parasites of bacteria, phage persistence in a microbial ecosystem is dependent on the presence of a suitable sensitive host. Phage infection is typically followed by one of two replication cycles, lytic or lysogenic (39) (**Figure 1A**). In both cases a phage virion binds to the host cell surface using a phage receptor-binding protein triggering the insertion of its genome into the host. For lytic phage, subsequent translation of the phage genetic material by the host cell results in the replication of the phage genome, assembly of phage particles and lysis of the host. This results in the release of new phage virions into the environment that can infect nearby hosts. Alternatively, lysogenic infection results in the replication of the phage genome within the host cell without the immediate synthesis of phage virions. The phage genome may integrate into that of the host where it exists as a prophage, replicating together with the host genome and thus persisting in resulting daughter cells. In the case of pseudolysogeny, the phage genome persists as an episome within the host cell, separate to the host genome. In order to ensure subsequent daughter cells contain phage genomes pseudolysogenic phage can use maintenance systems such as toxin-antitoxin (40, 41). However, cases of daughter cells lacking pseudolysogenic phage have also been reported (41). Following an induction event, the lysogenic phage will initiate the translation of its genome and subsequent production of phage virions leading to host lysis. Additionally, phage such as M13 undergo chronic non-lethal infection cycles, where newly produced virions exit the cell without lysis (42). However, little is known about the prevalence of these different lifecycles in the human gut.

Resistance and Counter-Resistance

Bacterial hosts employ a wide array of phage resistance mechanisms which have been comprehensively reviewed by Labrie et al. (43) and Rostøl et al. (44). To prevent phage adsorption, bacterial cells can differentially express or mutate cell surface receptors (**Figure 1B1**) (45, 46), S-layer proteins (47), or produce protective cell surface polysaccharides (48). Additionally, bacterial hosts can reduce the numbers of phage particles available to infect hosts by producing outer membrane vesicles (49). These bind and sequester phage particles, reducing their numbers in the environment and thereby the risk of infection. Should the phage successfully bind to the appropriate surface receptor, the fate of the host is not yet sealed as anti-phage resistance mechanisms extend to all steps of the phage infection cycle.

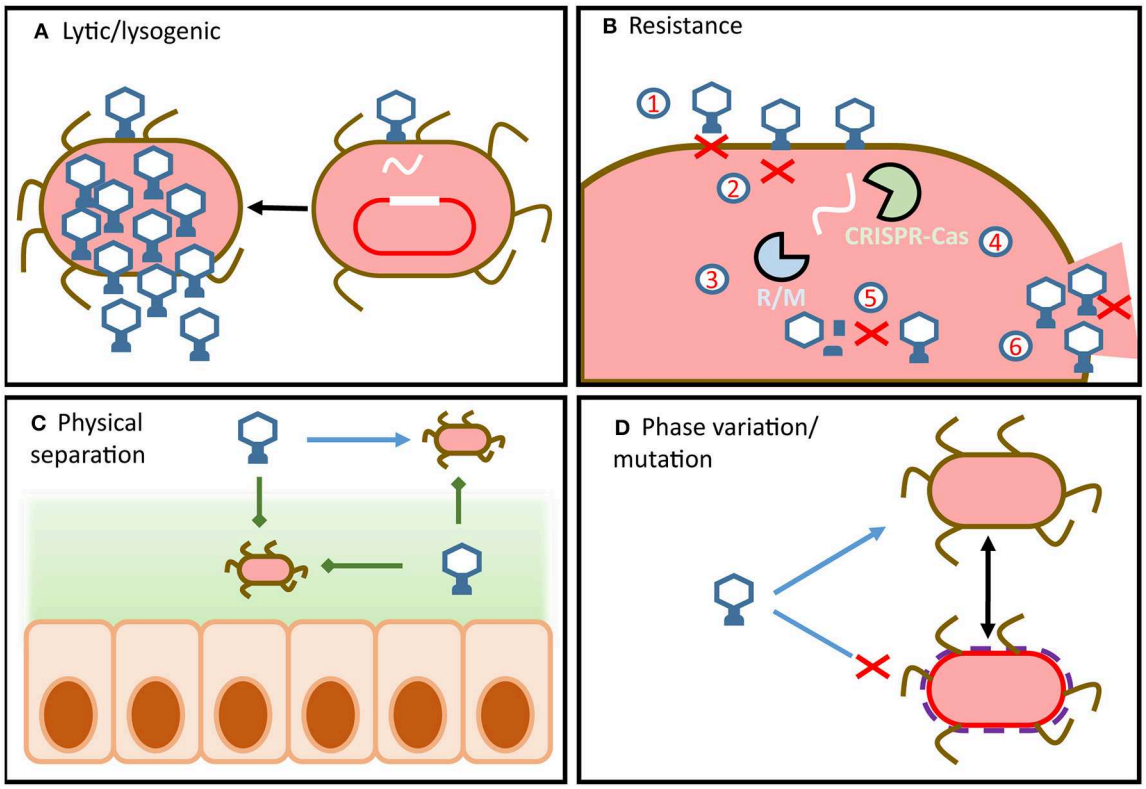


FIGURE 1 | Overview of phage-host dynamics in the gut. **(A)** Phage infection can lead to virulent or temperate replication cycles. Integrated temperate phage use internal and external signals from hosts to determine if or when to enter the lytic cycle. **(B)** Bacteria can possess a wide array of defense mechanisms which target different steps of the phage replication cycle. Similarly, phage encode a wide array of counter-defense mechanisms which target host defenses and allow the phage to remain infectious. **(C)** Physical separation of phage and host (e.g., in mucous or in lumen) means that dynamics change along the radial and longitudinal axes of the gut. **(D)** Strain-level variation can result from resistance by mutation or by phase variation.

Hosts can prevent phage DNA injection entirely by modifying inner-membrane proteins (**Figure 1B2**) (50), identify and degrade injected phage DNA using restriction modification systems (**Figure 1B3**) (51) and CRISPR-Cas systems (**Figure 1B4**) (44), chemically block phage DNA replication (52), or prevent virion assembly (**Figure 1B5**) (53).

Should these defense mechanisms fail to prevent phage replication within the cell, the bacterial host can sacrifice itself in order to protect its sister cells (54). These are referred to as abortive infection systems and act by shutting down cellular functions to prevent phage release (**Figure 1B6**).

Phage defense systems are regularly encoded on mobile genetic elements that can facilitate the transfer of resistance across the bacterial community. However, due to their metabolic costs bacterial cells rarely encode more than one of these systems (55). This gives rise to complex dynamics between hosts with the metabolic burden of resistance and susceptible hosts with fitness advantages. In addition, a cell carrying a prophage can be made resistant to other phage in what is referred to as superinfection exclusion (56). To further complicate these relationships, antagonistic coevolution of phage-host pairs has led to the development of phage counter-resistance mechanisms, which allow phage to remain infectious in the face of a resistant host population (57). Phage counter-resistance can range from glycosidases to degrade host capsules (58) and reveal binding sites, to directed mutagenesis or hypervariable receptor binding proteins (59–61). These systems allow phage to retain compatibility with modified host receptors and also allow for the expansion of host range. Phage can even overcome host CRISPR-Cas by mutating or deleting CRISPR target sites or expressing anti-CRISPR proteins to directly interfere with CRISPR-Cas activity (62). Some of the most striking phage counter-resistance mechanisms include alteration of phage DNA to evade host restriction modification mechanisms (63) and phage encoded CRISPR-Cas systems that target and disable a range of host defense systems (64).

Ecological Relevance of Phage-Host Dynamics

These mechanisms of infection, resistance and counter-resistance underpin virome-microbiome interactions. Thus, in order to investigate how the virome shapes or reflects the microbiome, we must first understand these interactions in the GIT. Understanding phage-host interactions is also vital if we are to use the virome as a diagnostic or therapeutic tool in the future. To this end, numerous ecological models have

been used to describe phage-host interactions in the context of a biological system. Some models focus entirely on the interplay of resistance and infectivity such as the arms race dynamics model (65). This model proposes that phage infection applies selective evolutionary pressure for mutations in the hosts, resulting in resistant host populations. These mutations in turn select for phage mutations that restore infectivity, resulting in predator prey cycles (66). Other models take into account phage-host density and the metabolic cost of resistance such as the fluctuating-selection dynamics model (67, 68). This proposes that as phage predation selects for resistant hosts, it will also reduce the number of phage virions in the environment. This absence allows for the expansion of susceptible bacterial strains which lack the metabolic burden of resistance and therefore out-compete resistant hosts in the absence of phage. This transient resistance and infectivity of phage-host communities results in short-term fluctuations of phage and host numbers, but the long-term persistence of both (69).

It is important to note that strain-level fluctuations of phage-host are difficult to study in the context of the microbiome, making it difficult to verify or quantify this model. Strain-level variation within bacterial hosts cannot be detected by 16S rDNA analysis (67) and hampers metagenomic assembly of phage (38, 70). Despite these challenges, recent insights have proposed that variation in capsular polysaccharides encoded by the abundant gut bacterium *Bacteroides thetaiotaomicron* play a central role in phage susceptibility (71). This mechanism supports the concept of fluctuating-selection dynamics, as phase variation of the capsular polysaccharides creates heterogeneous host phenotypes within an isogenic population. This in turn leads to transient phage resistance across the population as host phenotypes are dynamic and non-uniform (72). Additionally, this resistance can occur without the need for horizontally-transferred resistance or mutation (**Figure 1D**).

Phage replication not only requires the host to be present and susceptible but it must also be metabolically active. It has therefore been proposed that the metabolic state of the host is one of the primary barriers to phage infection (69). In this way bacterial hosts can be transiently resistant to phage infection due to a lack of nutrient availability and a dormant growth phase, without incurring the metabolic cost of encoding resistance (**Figure 1D**) (73). Additionally, this mechanism would support the proliferation of both phage and host in the GIT, independent of resistance/counter-resistance dynamics. Nutrient availability varies significantly along the GIT, which implies that phage-host dynamics in the proximal or distal colon may be significantly different to those in fecal samples (**Figure 1C**) (74, 75). As the majority of virome studies draw conclusions from fecal samples, our current understanding of phage host dynamics in the GIT is limited.

The kill-the-winner ecological model is an extension of lotka-volterra dynamics applied to phage-host interactions. This model describes rapid changes of diversity and abundance of both phage and their hosts. As the most abundant bacteria are killed by their phages, other bacterial taxa will take over the ecological niche and be subsequently killed by their phages. In this way high levels of phage-host diversity and abundance are maintained (76). In ecosystems where lotka-volterra or kill-the-winner dynamics can be applied, phage exhibit an exclusively predatory relationship on hosts and the microbial biomass is significantly below the carrying capacity of the ecosystem (77, 78). However, in the human gut ecosystem microbial biomass approaches the carrying capacity of the ecosystem and the virus to microbe ratio (VMR) is low (79). Despite reports of kill-the-winner dynamics in infants (80, 81), this suggests that these models cannot not fully explain phage-host interactions in the healthy adult gut (76). Additionally, these dynamics overlook lysogeny and conditions that govern the switch between lytic and lysogenic replication cycles. Furthermore, it has been proposed that the gut virome is dominated by temperate phage (82) and that compositional changes in temperate phage communities are associated with disease states (25). Consequently the mechanisms that determine the switch between lytic and lysogenic replication cycles are also central to understanding virome-microbiome dynamics. Recent reports have described phage which can hijack bacterial quorum sensing machinery to determine the density and metabolic activity of the bacterial population from within the host cell (83). This in turn, could dictate whether persisting within in the host genome or excising and entering the lytic cycle would favor phage proliferation (**Figure 1A**). In addition, single cell analysis of phage-host interactions between the temperate P22 phage and *Salmonella typhimurium* suggested that phage were directly involved in creating transiently resistant host subpopulations (41). This allowed for both lysogenic and lytic replication without impeding host proliferation. Despite these intriguing insights, little is known about the dynamics of temperate to lytic switching in the mammalian GIT, highlighting a crucial target for future virome research.

Ecological models that consider the switch between lytic and lysogenic replication such as the piggyback-the-winner model appear to support experimental evidence of phage-host dynamics within the mammalian GIT (84, 85). This model focuses on a lytic to lysogenic switch that is host density dependent. Traditionally phage were believed to enter the lysogenic cycle in cases of high VMR (i.e. increased phage abundance, decreased host abundance) as a means to persist in the environment until host density can support lytic replication cycles. However, experimental evidence from coral reef ecosystems suggested that phage also entered the lysogenic cycle in high host density situations (84). Subsequently, phage can "piggyback" on host success in the ecosystem at that particular time. This model has also been proposed for phage-host interactions on mucosal surfaces in the GIT. It was suggested that at the mucosal surface, high bacterial colonization and high VMR gives rise to piggyback the winner dynamics, whereas deeper mucosal layers may give rise to kill-the-winner dynamics due to the lower levels of bacterial colonization and low VMR. This model is also supported by reports that rapidly evolving Ig-like domains expressed on phage capsids interact with mammalian host mucus glycans. This in turn results in subdiffusive motion of phage within mucus and allows them to persist in mucosal layers of the gut (86, 87).

While these models assist in our understanding of how the virome interacts with the microbiome and the gut environment, it is also important to consider their limitations. Phage-host dynamics can be expected to change both radially and longitudinally within the GIT (88) to reflect physical separation of phages and hosts (**Figure 1C**) and metabolic changes in the host populations (**Figure 1D**). They will also be heavily influenced by dietary components and the composition of the fecal matrix itself (5). To this end, sampling method and sample composition must be considered when drawing conclusions from virome data (**Figure 2A**). In the absence of studies of how these dynamics change along the human GIT, phage-host interactions must be interpreted as a snapshot of one particular point in space and time.

COMPOSITION OF THE GUT VIROME

The gut virome consists of two elements, the temperate phage located within bacterial genomes and the free virions or virus-like particles (VLP). The VLP fraction is obtained by applying several physical and enzymatic steps that remove dietary debris and prokaryotic and eukaryotic genetic material (**Figure 2B**) (29). As VLPs represent only a small fraction of the mass of the microbiome, virome studies that do not carry out this viral enrichment are limited primarily to the prophage sequences within bacterial cells (89, 90). However, these prophage sequences are in turn under-represented in studies that focus on the VLP fraction. In order to understand the virome as a whole, both elements need to be analyzed in tandem. The majority virome studies focus on one of these elements in isolation and as a result the apparent structure and composition of a virome is heavily dependent on the physical preparation of virome samples in the laboratory, the sequencing strategy, and the bioinformatic methodology employed (**Figure 2**). This is further complicated by the limited representation of human gut viruses in databases and a reliance on phage taxonomic classification systems which do not necessarily represent viral biology. Consequently, our understanding of the taxonomic composition of the phage populations in the human gut are predictably varied and contradictory.

Taxonomic Composition

Sequence-based analysis suggests the identifiable fraction of the gut virome is dominated by small single-stranded DNA (ssDNA) phage of the *Microviridae* family and double-stranded DNA (dsDNA) phage of the order *Caudovirales* (26, 29, 30,

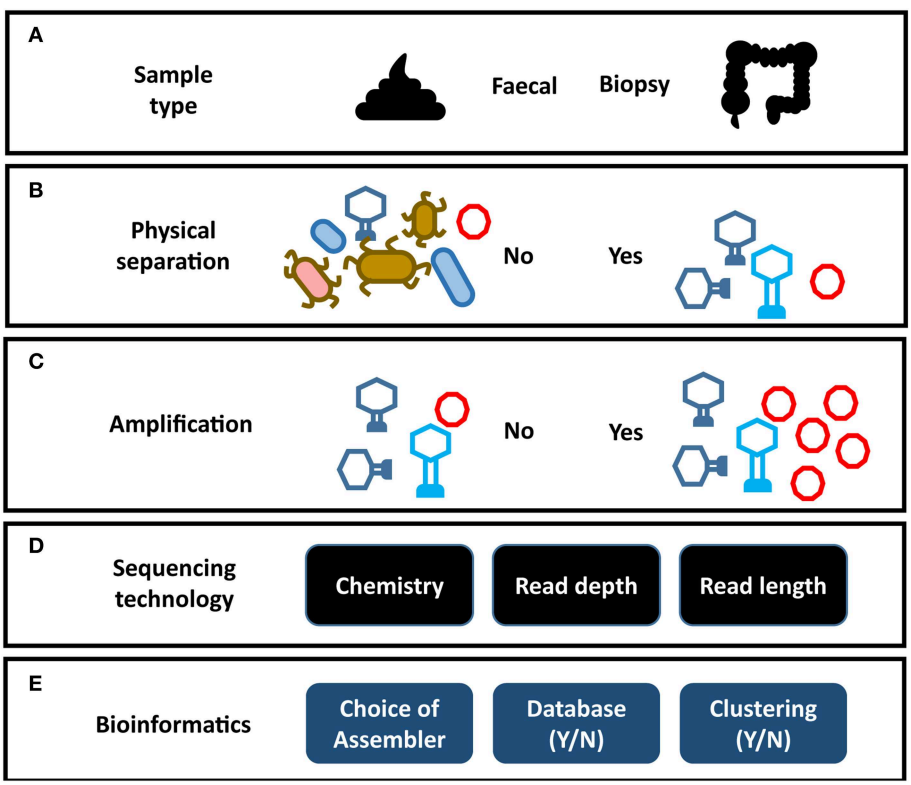


FIGURE 2 | Impact of analysis choices on virome composition. Each step of a virome analysis protocol presents different options, each of which may affect the final outcome. **(A)** Sample type. **(B)** Physical separation of VLPs. **(C)** Amplification of virome DNA can preferentially amplify certain viral taxa (see **Figure 3**). **(D)** Sequencing chemistry, depth of sequencing and read length. **(E)** Assembly programs vary significantly in their ability to assemble virome data (see **Figure 3**). Reporting on the composition of viral sequences with homology to reference databases excludes the unknown majority of the virome. Clustering viral sequences by gene composition offers a promising alternative to database dependent methods by addressing high levels of sequence divergence in viral genomes.

60, 91). However, bias in extraction methods has also been reported to skew the abundance of *Microviridae*, calling into question their dominance in the virome (discussed later) (92) (**Figure 2C**). The *Caudovirales* are classified into three families according to their distinctive virion morphology, consisting of a head and a tail structure making them easily distinguishable in microscopy studies. *Siphoviridae* exhibit long (sometimes up to 1 μm in length), non-contractile tails, *Podoviridae* exhibit short non-contractile tails, while *Myoviridae* exhibit a more rigid contractile tail composed of distinctive sheath proteins. These phage families have linear genomes, can encode a relatively large gene repertoire (in extreme cases containing over 600 genes) and can exhibit both temperate and lytic replication cycles. Early sequencing studies of the virome reported changes in *Caudovirales* community composition being associated with disease (25, 93, 94) and high *Caudovirales* abundance in EM images (28, 95). However, reported abundance in EM images could also be influenced by the fact that they are more easily identified than other viral taxa. Possibly as a result of these observations, *Caudovirales* composition and abundance has become a regular focus of virome studies, where it is used as a proxy for virome composition as a whole. However, as known *Caudovirales* tend to represent a small minority of VLP sequences in the human gut, caution should be taken when interpreting these results.

Due to the current structure-based classification system of *Caudovirales*, phage that exhibit significant functional similarities can be considered members of different families due to tail morphology. For example, phages P22 and lambda are classified as *Podoviridae* and *Siphoviridae* respectively, despite both undergoing the same replication cycle, sharing significant similarity in gene sequences and exhibiting identical genome organization (i.e., gene order and layout and regulation of transcription). Additionally this classification system is not grounded in sequence or protein sharing, yet sequence and protein homology are the primary methods of identifying *Caudovirales* in viromes (27, 34, 96, 97). This classification system is therefore limited in its ability to reflect the biological role or interactions of sequences classified as *Caudovirales*. A number of recent studies have also highlighted anomalies in *Caudovirales* taxonomy (98–100) and have proposed novel sequence-based methods to restructure the viral order (99). However, these are not without their own challenges, as shared genes and gene cassettes have been found to blur the boundaries between various dsDNA viruses (101, 102), which will be discussed in detail later.

Case Study of the Most Dominant Gut Phage, CrAss

A prime example of the limitations of focusing on identifiable *Caudovirales* alone when carrying out virome analysis is that of crAssphage, one of the most abundant and successful biological entities of the mammalian gut. First reported in 2014 using novel assembly methods (103, 104), crAss was found to be six times more abundant in publically available gut metagenome samples than all other phage together. It made up to 90% of VLP sequencing reads and 30% of whole metagenomes in certain individuals, yet did not share homology with known reference databases (104). As a result, crAss would have gone entirely unnoticed using database-dependent analysis alone. CrAss has since been found to be globally distributed, with strains reflecting geographic distribution of human populations (105). Intriguingly, the presence of distant relatives to crAss in primates may suggest that crAss has coevolved with humans for millions of years from ancestors shared with primates, to modern *Homo sapiens* (105). These results highlight the importance of the unknown majority of the virome in human health and why it is critical to analyze the virome as a whole. Given the proposed ubiquity and dominance of crAss in the human gut it is also possible that the abundance of *Caudovirales* reported by EM studies may in fact reflect members of the extended crass-like phage family, which at the time were unknown.

The characterization of crAss and studying its role in the gut microbiome has been hampered by a lack of database representation, unknown host range and until recently (106), unsuccessful attempts to form plaques on agar overlays. However, its progression from an unknown abundant phage sequence to a characterized dominant member of the gut virome provides a useful framework in building our understanding of viral dark matter. Its initial discovery was built upon by using sensitive protein homology searches to identify an extended family of crAss-like phage from human gut virome samples (107, 108). Although sequence similarity across family members was low, protein family-based clustering identified conserved capsid proteins and predicted crAss to encode a short tail similar to that of the *Podoviridae* family of *Caudovirales* (108). Subsequently, crAss has also been identified in patients with diarrhea and in Malawian infants (107). Additionally it was found to be shared across healthy individuals and was capable of stable engraftment in FMT treatments (109). However, the host range and mechanism behind crAss ubiquity and abundance remained unknown.

Through the use of enrichment-based techniques the host for one member of the crAss family was confirmed to be *Bacteroides intestinalis*, a finding previously suggested by co-abundance profiling (104, 106). This provided new opportunities to study the phage-host interactions of the most abundant phage family in the human gut and its role in the microbiome. The ability to culture crAss *in vitro* also led to the description of one of the most intriguing aspects of this phage, its ability to coexist with its host at high abundance. This has a profound impact on our understanding of how the virome shapes the microbiome, as it appears that in crAss-rich individuals the dominant phage populations do not restrict the growth and proliferation of their host. Furthermore, it is believed that crAss persistence and coexistence with its host is not due to crAss undergoing temperate replication, as it does not possess any of the genes typically associated with lysogeny (106). Additionally, crAss prophage sequences have never been observed in the numerous *Bacteroidales* genomes available in sequence databases. It is possible that crAss replicates in a pseudolysogenic manner, existing as an episome within the host cell (41), that it facilitates a chronic infection without killing the host cell (42), or that it can exist in an extracellular carrier state (110). The high proportion

of resistant host cells in culture (1–2%) could also suggest that a host cell carrying crAss in a pseudolysogenic state could confer resistance from infection as in superinfection exclusion. However, experimental evidence suggests crAss replicates in a lytic manner and its abundance in the gut is maintained by transiently resistant hosts. As phase variant CPS and transient resistance have been found to be central in phage-host dynamics in other *Bacteroides* species (71), it is possible that these mechanisms are also central to the unusual phage-host dynamics observed in crAss both *in vitro* and *in vivo*. This theory is supported by the reversion over time of some resistant clones to sensitive states (106). Given the abundance of *Bacteroides* and crAss in the human gut, phenotypic heterogeneity across a host populations may be the central mechanism to support stable interactions between some lytic phage and their hosts in the microbiome.

Another possible explanation for sustained crAss-host proliferation is the recently proposed "Royal Family" ecological model (111). This suggests continuous kill-the-winner dynamics occur at a strain-level rather than at a species or genus-level. In this way, the abundance of both phage and host would appear to be stable as fluctuations would occur below the level of detection. Subsequently, detailed analysis of strain-level variation between crAss and *B. intestinalis* could provide insights into the importance of this model in the GIT. A similar study carried out by De Sordi et al. (112) supports the "royal family model" and gave intriguing insights into its mechanics. This study described a point mutation in the tail fiber gene of P10 which is associated with strain-level host expansion from *Escherichia coli* LF82 to *E. coli* MG1655. Interestingly, this strain-level host switching was only observed when phage and both hosts coevolved within the murine model and was not observed when the two strains were cultured separately. Subsequent experiments revealed that switching required the presence of an intermediate host *E. coli* MEc1. These observations suggest that crAss-host stability in the gut could be caused by a wide array of strain-level and phenotypic variation. However, as strain-level variation is difficult to observe *in vivo* and *in vitro*, novel analytical approaches may needed to reveal crAss phage-host dynamics in the GIT.

The proliferation of both phage and host as is seen in crAss-*B. intestinalis* dynamics could also suggest that the presence of phage could confer an ecological advantage to the host. This phenomenon has been regularly reported in the ocean, where phage infecting cyanobacteria were found to carry auxiliary metabolic genes encoding photosynthetic genes (22, 113). Similarly, phage-mediated transformation of the host has also been well-established in disease such as the lysogenic phage encoding shiga toxin (23) and cholera toxin (114). It is therefore highly likely that in the dense and diverse ecosystem of the gut, extensive horizontal transfer of genes between hosts is facilitated by phage infection. In this way the virome may play a crucial role in shaping the functional capacity of the microbiome. One report of the presence of significant numbers of antibiotic resistance genes in gut virome sequences (115) was later shown to have probably resulted from bacterial contamination and confirmed that examples of phage-encoded antibiotic resistance genes were rare (116). This could be due to the efficiency of phage replication and the fitness cost of carrying antibiotic resistance genes. Without the selective pressure of antibiotics, viruses that pay the metabolic cost of carrying antibiotic resistance genes could be outcompeted (116). Additionally, the selective evolutionary pressure of remaining infectious in the face of a constantly adapting host may outweigh the need to preserve the host from an antibiotic. The follow up study also highlighted the importance of using stringent alignment criteria and validating results when classifying sequences or proteins. Due to the extensive unknown fraction of the virome and the difficulty in benchmarking classification criteria, lenient cut-offs can often lead to false conclusions of virome composition and function (116–118).

It is tempting to propose that the piggyback-the-winner hypothesis could be extended by considering the possible fitness advantages of carrying a prophage over and above superinfection exclusion. Such advantages could include carrying a virulence factor or providing access to a novel nutrient source (119, 120). Prophage-encoded fitness advantages have been observed in a number of pathogenic bacteria. These include prophage-encoded toxins (23, 114), alteration of O antigens in *Salmonella* and *Shigella* (121, 122) and phage-encoded glycosyl transferase operons which drive *Salmonella* LPS diversity (123). Furthermore, the horizontal transfer of virulence factors through temperate phage was found to be increased in gut inflammation (124). Additionally, recent observations in *Staphylococcus aureus* prophage (125) have described the packaging of chromosomal host DNA in phage capsids through a mechanism deemed lateral transduction. This mechanism suggests that phage-mediated horizontal gene transfer occurs at much higher rates than previously thought and that it plays a role in disease. However, the extent to which lateral transduction mediates gene transfer in the gut microbiome remains unknown. Examining phage-encoded auxiliary metabolic genes and how they shape the functional capacity of the gut microbiome is hindered at a large-scale metagenomic level by the complexity of fecal samples themselves. Dietary components and the sheer abundance and diversity of bacterial cells in feces make it difficult to completely remove bacterial sequences from virome samples. Use of density gradients such as CsCl are reasonably effective at generating viral particles devoid of cellular contamination, but will introduce bias in favor of particular viral capsid types (126) and are not suited to large-scale projects due to their associated manual workload. As a result, background contamination exists in the vast majority of virome samples (29, 116, 117) making it difficult to determine if a gene is of viral or bacterial origin. This may be further complicated by potential gene transfer agents (GTAs) (127) in gut VLP samples. GTAs are defective phage virions that exclusively carry fragments of bacterial chromosomal DNA and are used by bacteria as a means of HGT. As a result sequences within GTAs are difficult to differentiate from background contamination. However, the presence and prevalence of GTAs in human gut microbiome remains to be seen. Overall, despite the evidence of phage shaping the functional capacity of host communities, it is challenging to determine the extent to which phage transfer genes relevant to human health within the microbiome.

Controversy Surrounding the Core Virome

The widespread geographical distribution and stability of crAss-like phage supports the concept of a core human virome, which was initially proposed by Manrique et al. (26). This was in response to a growing body of evidence that a core microbiome played an important role in human health. The study proposed a core of 23 viral sequences, one of which being the original crAss genome, which were shared across more than 50% of samples from an independent cohort of 62 healthy individuals. However, these findings were in stark contrast to the well-established belief that the human gut virome is highly individual-specific at a sequence level (29, 34, 35, 128). This disparity is largely due to the criteria used to define the presence of a viral sequence in a sample. If a single sequencing read from an individual could align to a particular viral assembly, the assembly was deemed to be present. This lenient criteria does not account for the modular nature and extensive gene sharing that occurs across dsDNA viral genomes (99, 101, 129). Thus, it would not be possible to differentiate the true presence of a viral sequence in a sample from the presence of a shared gene between two unrelated phage. This in turn, would lead to an inflated number of viral sequences being shared across individuals. However, the concept of a core virome has received support from a recent study with adult monozygotic twins, in which 18 contigs were found to be present in all individuals ($n = 42$) (34). Here, more stringent read recruitment criteria were applied to differentiate shared genes from the true presence of a viral sequence in a sample. Interestingly more than half of the viral assemblies identified across all individuals were homologous to crAss. It should be noted however, that these assemblies may also represent fragments of the same phage genome or family.

In contrast to these findings and to the proposed global distribution of crAss phage in human populations, the compilation of a large-scale gut virome database called into question the existence of a core human gut virome at a sequence level (24). By examining VLP metagenomes from 572 individuals, this study proposed that a core human gut virome does not exist. Recently, Shkoporov et al. (128) too made observations which support these findings by examining the virome of 10 individuals across a 12-month period. Here, a personal persistent virome (PPV) was observed that was composed of viral sequences detected in at least six of the 12 monthly time points. In accordance with previous longitudinal studies (60, 82), some viral sequences were present at nearly all time points within an individual. However, the virome was also highly individual-specific and viral sequences were not shared across the PPV of all individuals which supports the findings of Gregory et al. (24). This high level of inter-individuality in the gut virome hampers our understanding of the virome in disease as it is difficult to detect common viral signals within or between cohorts. While it is likely that the individuality of the virome is driven by infection and resistance dynamics, the level of taxonomic resolution at which the virome is studied is also a contributory factor (35). Sequence-based virome studies are carried out at the level of metagenomic assembly due to the absence of universal marker genes, limited database representation and established taxonomic organization. This represents species or strain-level resolution and is in contrast to the majority of bacterial metagenomics studies, which tend to be analyzed at higher taxonomic ranks such as genus or family. It is possible to find patterns across virome cohorts using a minor subset of known viral sequences and by excluding unknown sequences (25, 27, 130, 131). However, it is not known if these subsets represent the dynamics of the virome as a whole. To this end, a number of clustering programs have been developed that group viral sequences based on shared protein families (99, 129) such as vConTACT (99). This is a similar approach to that which was used to establish the extended crAss family despite low levels of nucleotide similarity between family members (107, 108). In this way, protein-based clustering of the virome can reveal compositional patterns across individuals that were not visible at a nucleotide level. Furthermore, this approach allows for both the known and unknown components of the virome to be included in analysis giving new perspectives to the virome in health and disease.

By applying this protein family clustering approach (102) to the same longitudinal cohort Shkoporov et al. (128) gained new insights into the existence and composition of a core gut virome across individuals. This phylogenetic core was composed primarily of crAss and *Microviridae* and was not identifiable at a nucleotide level. Intriguingly, this core was not composed of temperate phage which is in contrast to observations in previous reports (82). Furthermore, temperate phage were found to make up a minor subset of the core virome both within and between individuals. These findings suggest that mechanisms other than lysogenic replication are responsible for long term stability of the virome within healthy individuals. Moreover, they are in accordance with the global distribution and persistence of crAss in the human gut and ecological models such as the "royal family model." Upon clustering the stable fraction of individual viromes (PPV) the largest and most interconnected viral cluster as associated with known *Caudovirales* sequences. This is in accordance with previous observations (99, 129) and reflects extensive shared genetic content across this order. It is also likely that this extensive gene sharing influenced previous database-dependent reports of temperate phage and *Caudovirales* dominance in the human gut virome. Furthermore, it highlights the importance of considering shared genes and gene cassettes when setting alignment criteria.

GUT VIROME IN DISEASE

Given the extensive evidence that phage can shape the composition and function of bacterial communities, the virome of the human gut has been studied in a number of diseases. However, as with the concept of the core virome, findings have been somewhat contradictory and any potential role for the virome in shaping the microbiome in disease remains elusive. Studies have reported gut phage populations were not significantly altered in diseases such as colorectal cancer and HIV-associated AIDS (130, 132), despite established associations between the gut microbiome and these diseases (15, 133). This contradicts the established view that the gut virome and gut

microbiome are closely linked and is more likely to reflect limitations of different analysis methods. These limitations include lenient alignment criteria to reference databases and the exclusion of viral dark matter from analysis. Furthermore, reports of changes in phage populations associated with diseases are often limited to changes in the composition of known *Caudovirales* (25, 90, 97, 131). Given the limitations of *Caudovirales* taxonomy and the challenges presented by extensive gene sharing across the order, these findings provide little insight into any role of the virome in disease.

An intervention study by Gogokhia et al. (134) sought to target cancer associated bacteria adherent invasive *E. coli* and *Fusobacterium nucleatum* with lytic phage in a germ-free mouse model. However, a direct interaction between the mammalian immune system and phage virions resulted in an exacerbated colitic reaction. The authors also proposed that phage DNA plays a central role in phage interaction with the mammalian immune system as empty phage capsids did not induce an immune response. Similarly an *in vitro* study using *S. aureus* and *Pseudomonas aeruginosa* phage observed a production of both pro and anti-inflammatory cytokines from peripheral blood mononuclear cells following endocytosis of purified phage virions (135). These observations are supported by the proposed ability of phage virions to cross the mammalian epithelial barrier *in vitro* via peptide sequences expressed on the capsid surface (136, 137). In this way it is possible that phage communities in the human gut shape the gut microbiome indirectly through interactions with the mammalian immune system. This concept of phage translocation and interaction with mammalian immune system has also been discussed in a number of reviews and perspective pieces as follows (138–141). Through the induction of a pro or anti-inflammatory response, phage could facilitate conditions that would favor a particular host or replication cycle. It is also possible that proposed phage–immune system interactions are driven by bacterial populations to facilitate infection or persistence in the human body. This was first demonstrated by lysogenic Pf phage which triggered maladaptive viral pattern recognition receptors and facilitated the chronic infection of *Pseudomonas aeruginosa* in murine and human cells (142). This was also the first reported case of a directly pathogenic effect of phage in bacterial infection and demonstrated that phage do not need to directly encode virulence factors to impact the virulence of their host.

Fecal Microbiota Transplantation

FMT (fecal microbiota transplantation) is an emerging and experimental therapy that aims to restore healthy gut function through infusion of a fecal slurry from a healthy individual to the colon, cecum, or duodenum of the recipient. It has been shown to be very effective in the treatment of recurrent CDI with donor bacteria colonizing recipients for up to a year (143) and reported success rates of 80–90% (144). The first evidence that the virome had potential as a tool to shape the microbiome and may play a role in the efficacy of FMT treatment was reported by Ott et al. (145). In this study, patients with relapsing CDI received fecal filtrates from healthy donors that resulted in CDI symptoms being eliminated for up to 6 months. Furthermore, recipient phage populations were substantially altered, resembling those of the donor for a minimum of 6 weeks. Surprisingly, *Lactococcus* phage were reported to dominate both donor and recipient virome, despite *Lactococcus* spp. representing only a minor fraction of the gut microbiome. This could reflect a dominance of lactococcal phage in the donor and recipient, implying lactococcal phage play an important role in homeostasis in CDI. However, phage sequence databases are dominated by those which are industrially relevant and cultivable, which includes lactococcal phage (146). As a result, unknown sequences are statistically more likely to align to lactococcal phage and other cultivable or industrially relevant sequences when lenient alignment criteria are used. Thus, the dominance of lactococcal phage in the virome of FFT recipients is likely to be yet another artifact of database-dependent analysis methods. It should also be noted that, in accordance with the majority of database dependent virome studies, lactococcal phage are members of the order *Caudovirales*. This supports the concept that *Caudovirales* dominance reported at an order and family level in the gut virome could be driven by gene modules shared with *Caudovirales* in reference databases. However, the resemblance of the donor virome to that of the recipient suggests that, regardless of classification limitations, the virome plays a significant role in the maintenance of homeostasis in the gut. A subsequent study by Draper et al. (109) also reported the stable engraftment of donor phage populations in recipients for up to 12 months. In accordance with observations at a bacterial level (143, 147) successful phage engraftment was also dependent on specific donor-recipient pairings. It should be noted that the role of other elements present in the filtrate (i.e., chromosomal DNA, plasmids, bacterial cellular components, and signaling molecules) remains unknown and could also play a role in the restoration of healthy gut function in CDI. With conditions such as ulcerative colitis (UC) that are characterized by more subtle microbiome changes than CDI, successful FMT treatment (i.e., remission and mucosal healing) was not associated with changes in the phage population (148). This is in contrast to reports of alterations in bacterial diversity following FMT in UC (149). Additionally, changes in the diversity of phage populations between healthy and UC cohorts were not observed (148). These observations were in contrast with previous IBD virome studies, which reported differences in phage alpha diversity between healthy and UC cohorts (25, 97).

Inflammatory Bowel Disease

Inflammatory bowel disease is a prevalent chronic disorder of the gastrointestinal tract with both genetic and environmental risk factors (150). The composition of gut bacteria and their interaction with the host immune system are believed to be central to its pathology, yet the etiology of the disease remains poorly understood. Given the evidence that the virome can interact directly with the host immune system and shape the composition and function of the microbiome, both fecal and mucosal phage communities have been studied in IBD. Current understanding of phage populations in IBD has focused on VLPs, proposing that disease-specific patterns of *Caudovirales*

are linked to Crohn's disease (CD) and ulcerative colitis (UC). Furthermore, these changes in VLP composition have been reported not to reflect alterations of the bacterial community. However, the details of these compositional changes vary between studies (25, 27, 93–95, 151).

Early IBD virome studies using the Roche 454 sequencing platform were limited by sequencing depth. These studies observed lower diversity and greater variation in the fecal VLP communities of patients with CD relative to healthy samples (93). The same group performed another sequence based analysis of the fecal and mucosal VLP community of CD and observed greater viral load and diversity in the feces than mucosa of all individuals (151). Additionally, virome alpha diversity was reported to be significantly lower in disease. However, in contrast to their previous study, it was also reported that both healthy and CD cohorts were dominated by *Microviridae* rather than *Caudovirales*. Another study that analyzed the first mucosal virome in pediatric CD (94) proposed a dominance of *Caudovirales* phage overall and detected a single viral sequence in colonic mucosal samples from patients with CD. While this may suggest an extreme dominance of this virus in the mucosa of pediatric CD it more likely reflects insufficient sequencing depth resulting from low biomass samples and should be treated with caution.

As sequencing technology progressed, researchers were granted more detailed insights into the virome in IBD. However, these insights contradicted previous findings and highlight the impact of sequencing platform on results. This could also suggest that our understanding of the virome in disease will continue to change as sequencing technology progresses. Illumina-based studies reported disease-specific increases in *Caudovirales* alpha diversity in the VLP viromes of UC and CD compared to healthy controls in adults (25) and children (97). Intriguingly, these alterations were also reported not to reflect changes in the bacterial community (bacteriome). Conversely, decreased alpha diversity was observed in *Caudovirales* families from the mucosal VLP virome of UC relative to healthy controls (27). This supports the idea that viral communities and ecological models differ at different spatial locations within the gut. Additionally, it suggests that phage of the order *Caudovirales* play a central role shaping the microbiome in IBD. However, as with other studies of the virome in disease, these findings are limited to minor identifiable fractions of the dataset and offer little by way of insight into the role of phage in IBD.

In an effort to expand our understanding of the gut virome in IBD and move beyond the limitations of viral sequence databases, Clooney et al. applied the whole-virome analysis (WVA) (35) protocol discussed earlier to the keystone IBD virome dataset published by Norman et al. (25, 35). In this way, it was possible to gain novel insights into the composition viral dark matter in this disease. Contrary to the original findings of the study, alterations in the whole virome mirrored those of the bacteriome, and differences in overall virome alpha diversity were not seen across the whole-virome. In accordance with current understanding the gut virome, high levels of inter-individuality were observed, and were likely to conceal any patterns in virome composition across individuals. Subsequently, we followed the protocol established in Shkoporov et al. (128) to cluster viral sequences according to gene content. This revealed a core of primarily lytic phage in healthy individuals and supported observations of Shkoporov et al. (128). However, this healthy core was also found to be absent in patients with IBD where it appeared to be replaced by a community of temperate phage. The majority (six) of the eight viral clusters which made up the healthy core virome in this analysis did not share homology to known viral sequences further highlighting the biological signals that may be missed when relying on database-dependent methods. Interestingly, one of these healthy core virome clusters was identified as crAss. This supports previous observations of its ubiquity in healthy human populations, low rates of detection in unhealthy individuals and its role in the core healthy human virome.

One possibility is that in the inflamed gut, environmental stresses from the human immune response such as antibodies and reactive oxygen species, leads to increased induction of the prophage present in the bacterial microbiome. The physiology of bacterial cells and the composition of the bacterial community influences whether integrated prophage enter the lytic or lysogenic replication cycle (**Figure 1A**) (83, 152). It is therefore likely that temperate virome would react to environmental stress and resulting changes in the host community. This increased lytic replication and subsequent death of bacterial hosts would correspond with the observed reduction in bacterial alpha diversity associated with IBD. It would also correspond with the observed increase in in free phage virions (95) and temperate VLPs in disease (35). Additionally, the resulting increase of bacterial cell wall components and debris available to interact with the human immune system could perpetuate an inflammatory response. In accordance with *in vitro* reports, it is also possible that increased cell wall permeability associated with inflammation allows for increased phage translocation (137) and interaction with the host immune system (134, 135, 153). The observations of Clooney et al. (35) provide a novel theoretical, mechanistic rationale for the interaction of the whole virome and the microbiome in disease, beyond taxonomic assignment and compositional patterns of the known minority. Additionally, they provide comprehensive evidence to support the mechanisms that had been previously proposed by Norman et al. (25).

The extent to which the switch from temperate to lytic replication cycles and the composition of the core virome shape the gut microbiome and influence human health and disease remains speculative. However, the analysis approach outlined by Clooney et al. (35) paves the way to a better understanding of how the interplay between the microbiome and the virome reflects or influences human health and disease. By allowing the detection of biological signals across the entire virome it is now possible to identify viral signals associated with disease which had been previously undetected. The WVA approach is also supported by the methods used to characterize the crAss-like family (106–108). CrAss has progressed from an unknown viral sequencing anomaly to providing insights into the composition and function of the healthy human gut virome. In

turn, the methods used in this progression provide a framework to characterize unknown but biologically relevant sequences identified by WVA.

It should be noted, that although many virome studies tend to report the dominance and composition of known *Caudovirales*, this provides little insight into the biological impact of these phages or how they shape the gut ecosystem. *Caudovirales* dominate reference databases, exhibit extensive gene sharing across families and orders and feature temperate phage genera (101, 102, 154). It is therefore likely that database-dependent methods of virome analysis classify unknown dsDNA virus sequences as known *Caudovirales* due to shared genes or gene cassettes and lenient detection criteria. As the representation of gut virome sequences in databases improves with efforts such as those carried out by Gregory et al. (24) and the characterization of crAssphage (106, 107), database-dependent analysis methods may be able to better reflect the composition and dynamics of the virome. However, as extensive gene sharing remains a central part of dsDNA viral genomes it is essential that stringent alignment criteria are used to differentiate shared functional modules from the presence of a particular virus (36, 155), regardless of the database used. Additionally, it is crucial that these alignment methods and their findings are validated to avoid misleading conclusions as to virome composition or function (116–118, 145).

ADDRESSING THE CURRENT CHALLENGES OF VIROME RESEARCH

As has been discussed, studying phage of the human gut microbiome is made challenging by the composition of the sample (usually a fecal specimen; **Figure 2A**). In order to enrich for the VLP fraction of the virome, extensive mechanical, chemical, and enzymatic processing is required to remove cellular DNA and dietary components (**Figure 2B**) (29, 126). Unfortunately, this results in particularly low DNA yields that can complicate the generation of sequencing libraries. This challenge is more pronounced in mucosal virome studies where DNA yields are lower again (132, 156). As a result, all but one (26) virome study to date have depended on multiple displacement amplification (MDA) of viral DNA to reach sufficient quantities to sequence. As with all metagenomics, it is crucial to find the balance between sequencing chemistry, depth of sequencing and read length. These factors have profound impacts on the final virome sequences available for downstream analysis. This was highlighted by the differences in virome alpha diversity reported by 454 pyrosequencing, when compared to deeper sequencing on the Illumina platform as previously discussed. Short read platforms such as the Illumina HiSeq are a means to perform deep sequencing of the virome with low error rates and relatively low input DNA requirements. However, these libraries can also lead to fragmented assemblies and poor recovery of viral genomes (38). Long read sequencing offers a promising solution to this assembly challenge, as it is possible to sequence entire viral genomes on a single read (**Supplementary Table 1**). This overcomes hypervariable sequences and repeat regions in viral genomes which hamper assembly (61). Currently long read sequencing platforms also require very large quantities of unfragmented DNA, which can be challenging acquire from virome samples. As a result, the initial DNA yield and the amplification step directly influence the sequencing chemistry, read depth and read length which can be used with virome sequences (**Figure 2D**). The MDA step has also been reported to introduce considerable bias into the composition of the resulting virome which must be considered when drawing conclusions from data (**Figure 3A**) (157, 158). Studies have also reported MDA preferentially amplifies small circular ssDNA viruses, which include the family *Microviridae* (**Figure 2C**) (92). This could call into question both the reported abundance and ubiquity of this family across individuals. Although it is difficult to quantify the extent to which preferential amplification occurs, recent meta-analysis of gut virome studies suggested *Microviridae* may be 10-fold lower in abundance than previously thought (24). It is believed that priming biases of the random hexamers used in the MDA reaction do not prime equally across all genomes, making quantitative interpretation of virome data difficult.

MDA bias can also have a significant impact on qualitative analysis of the gut virome. As the concentration of the DNA template also impacts the products of an MDA reaction, initial log-fold differences in the abundance of viral sequences are exaggerated by MDA (159, 160). This results in extremes of both high and low read coverage and uneven representation of the initial metagenome. As high-abundance sequences sequester sequencing resources, low-abundance sequences can be insufficiently covered (161). These coverage extremes have profound impacts on a number of steps in bioinformatic pipelines, but in particular metagenomic assembly (38, 161). A recent comparison of all assemblers used in virome studies to date observed that both high and low-coverage sequences resulted in fragmented assemblies and recovered only small proportions of viral genomes (38). Furthermore, samples were spiked with an abundant (10^7 plaque-forming units ml^{-1}) exogenous lactococcal phage Q33. These samples underwent identical extraction amplification and sequencing and resulting viromes were then assembled using all assembly methods which had been reported in virome studies at the time (16). Ten of the assemblers either failed to recover or significantly fragmented the Q33 genome and only one assembler recovered the genome at the correct length (**Figure 3B**) (38). These results suggest that numerous and potentially biologically relevant viral sequences may be not only be skewed in abundance but also excluded by current virome analysis protocols. This also means that we see the virome through the lens of the extraction protocol before any decision has been made to use database-dependent or independent methods (**Figure 2**, **Supplementary Table 1**).

As with many microbiome studies, conclusions from virome studies are primarily drawn from relative rather than absolute abundances of sequences. As has been discussed, these abundances are often skewed by MDA bias. This ambiguity highlights the need for quantitative analysis protocols

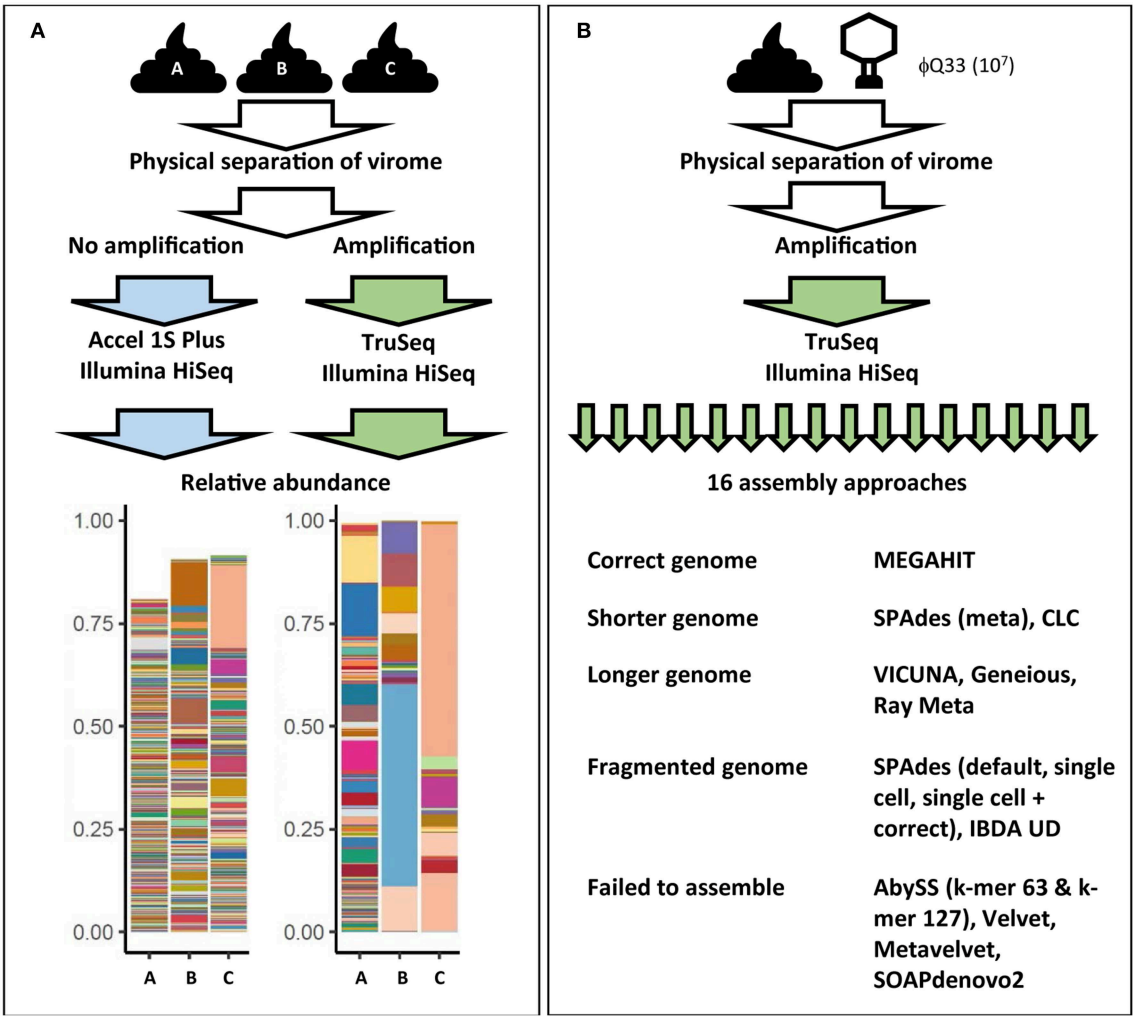


FIGURE 3 | Examples of how virome composition is influenced by key steps in analysis. **(A)** Three samples were subjected to identical filtration and DNA extraction steps. One set was amplified and prepared for sequencing using the Illumina TruSeq library kit while another set of unamplified samples were prepared using the Accel 1S Plus kit. Both sets were sequenced on the Illumina HiSeq platform. Differently treated samples differ in terms of final composition, represented in bar plots. Each color represents the relative abundance of a unique viral contig in each sample. Abundance does not reach 100% in the unamplified sample as the higher level of richness also hampered assembly [adapted from (128)]. **(B)** Impact of assembly software on final virome composition. A fecal samples was spiked with φQ33 10^7 PFU ml^{-1}, extracted and sequenced. These sequences were assembled using 16 assembly programs. Only one assembler identified the genome in a single contig of the correct length. Five assemblers completely failed to assemble the genome and a further five generated fragmented assemblies (38).

in virome studies as was recently described (29, 35, 128). These studies reported significant differences in the overall viral load between individuals. This total viral load was also correlated negatively with viral alpha diversity within the sample and the presence of abundant and typically virulent phage such as *Microviridae* and crAss. These results suggest that high viral load is associated with a low number of abundant phage, which consequently mask underlying temperate phage diversity. This also suggests the maintenance of high-abundance non-temperate phage may be closely linked to the health status of the gut. Subsequently, this may act as a useful biomarker for disease and provide insight into the phage-host dynamics related to microbiome stability and disease status.

FUTURE PROSPECTS AND CONCLUSIONS

Sequence-based analysis of the bacteriophage in the human gut has revolutionized our view of the gut virome and its relationship with the microbiome. However, this new insight has also revealed how little we know about this relationship. Our current understanding is founded predominantly on extending knowledge generated in reductionist phage-host studies *in vitro*, or by large scale metagenomic studies of the VLP fraction. While these *in vitro* studies give detailed insights into the mechanics of individual phage-host interactions, the prevalence of these interactions in the gut ecosystem remains speculative. Additionally, numerous studies have also

reported that these interactions can change dramatically in the gut (112, 134). Large-scale metagenomic studies suggest that the virome and the microbiome are closely linked, but these studies tend to give broad overviews of subsets of the virome and lack details on the phage-host interactions. This leads to significant gaps in our understanding of how phage-host dynamics *in vitro* differ from phage-host dynamics occurring in the gut.

Application of similar analysis methods across studies (i.e., MDA, alignment to reference databases, reports on *Caudovirales* alpha diversity) allows for comparison across samples and studies. However, many virome studies present inconclusive or contradictory results that hinder the progression of the field. Arguably the overreliance on these analysis methods is largely to blame for the gaps in our understanding of the virome. This is also supported by recent observations that methodology had greater impact on the conclusions drawn from virome studies than health or disease status (24). As conclusions are drawn from minor fractions of data and as detection criteria do not take into account phage biology and evolutionary history, we must pose the question: is the gut virome field built on unstable foundations? With this in mind, due caution must be used when interpreting the findings of virome data. As virome data is particularly sensitive to methodological bias, conclusions must be considered in the context of the analysis methods used (**Figure 2**, **Supplemetary Table 1**). These limitations highlight the need for radically new approaches to studying the virome if we are to understand its role in shaping the microbiome in health and disease.

Developments in sequencing library kits such as the Swift Biosciences Accel 1S Plus kit or extraction protocols like the linker amplified displacement LADs (36), offer potential solutions to creating unbiased sequencing libraries from low-input DNA yields. Through the removal of MDA bias and spiking known concentrations of exogenous phage (29), it may be possible to gain new insights into the true composition and absolute abundance of the virome. As has been discussed at length, the sensitivity of virome data to methodological bias highlights the critical need for extensive optimization and validation of all steps of virome analysis protocols, from wet-lab extraction protocols (29, 36, 155) to bioinformatic pipelines (37, 38, 161, 162). However, standardization and consistency must not be at the cost of developing new methods. Furthermore, when characterizing viral sequences it is crucial to use stringent detection criteria to minimize the impact of spurious alignments and the influence of gene sharing across dsDNA viruses.

Significant progress has been made in increasing the representation of gut phage in reference databases and there is a growing consensus that viral taxonomy will soon move toward sequence-based taxonomy (24, 163–166). However, proposed protocols to add metagenomic sequences to current taxonomic systems have not yet been accepted (167). Given the dominance of unknown sequences in virome data it is therefore crucial to accept the current limitations of phage taxonomy. Rather than force the virome to fit current taxonomic systems, we propose that future studies should allow the virome to reveal its own targets for downstream characterization. Subsequently, we have highlighted a method to analyze the virome in its entirety using the WVA protocol. Furthermore, we have described a framework to characterize unknown but biologically relevant viral sequences that may be identified using WVA. In this way it may be possible to address the gaps in our understanding of phage-host dynamics in the human gut, and see existing datasets in a new light.

To what extent the phage of the human gut shape the microbiome will dictate whether it will be possible to use phage as a therapeutic tool in the future. There are significant gaps in our understanding of phage-host interactions which need to be addressed before we can reach any conclusions on the usefulness of phage as a biotherapeutic. By increasing our understanding of phage-host interactions in the gut, it may be possible to pave the way for therapeutic applications of phage in the human body. However, the limited insights we have been granted to date of phage-host interactions have also highlighted some significant hurdles facing phage therapy. Early evidence that the virome could play a role in the success of FMT (109, 145) suggests there may be a future in using the virome to shape the microbiome in disease. To date, the majority of phage intervention studies are based on single phage-host pairs (or cocktails containing limited numbers of phage) *in vitro* which have been shown to be significantly different to phage-host dynamics *in vivo* (112). For example, phage have been found to switch hosts and interact directly with mammalian immune cells *in vivo*, which has serious implications for the future of phage therapy (134, 153). Additionally the phage-mediated transfer of host virulence factors (23, 114, 168, 169) as well as direct pathogenesis of phage capsids (142) suggests phage could be a potential risk in therapeutic settings. These challenges are confounded with regulatory issues (170) and additional gaps in our understanding of the pharmacodynamics of phage in mammalian tissue. Having the potential to directly interact with the immune system (86, 87, 136, 137) given their larger size relative to other biological therapeutic agents makes phage a more complex therapeutic agent than any that have preceded them. However, in light of the increasing incidence of bacterial pathogens which are resistant to antibiotics, and given the promising results of some existing phage therapy trials (171–174), overcoming these challenges is critically important. Similarly the gaps in our understanding of how phage shape bacterial communities will need to be addressed if phages are to have a role in avoiding future global health issues.

AUTHOR CONTRIBUTIONS

TS and CH drafted and approve of this manuscript.

FUNDING

This research was conducted with the financial support of Science Foundation Ireland (SFI) under Grant Number SFI/12/RC/2273 a Science Foundation Ireland's Spokes

Programme which is co-funded under the European Regional Development Fund under Grant Number SFI/14/SP APC/B3032, and a research grant from Janssen Biotech, Inc.

ACKNOWLEDGMENTS

The authors thank Dr. Andrey Shkoporov, Dr. Lorraine Draper, Dr. Hugh Harris, and Dr. Adam Clooney for their critical remarks during the preparation of this manuscript. The authors also thank Dr. Andrey Shkoporov for providing data featured in **Figure 3**.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2019.00784/full#supplementary-material

Table S1 | Pros and Cons of key steps in virome analysis pipelines.

REFERENCES

1. Sender R, Fuchs S, Milo R. Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol.* (2016) 14:e1002533. doi: 10.1371/journal.pbio.1002533
2. Duncan SH, Louis P, Thomson JM, Flint HJ. The role of pH in determining the species composition of the human colonic microbiota. *Environ Microbiol.* (2009) 11:2112–22. doi: 10.1111/j.1462-2920.2009.01931.x
3. Espey MG. Role of oxygen gradients in shaping redox relationships between the human intestine and its microbiota. *Free Rad Biol Med.* (2013) 55:130–40. doi: 10.1016/j.freeradbiomed.2012.10.554
4. Ridlon JM, Kang DJ, Hylemon PB, Bajaj JS. Bile acids and the gut microbiome. *Curr Opin Gastroenterol.* (2014) 30:332. doi: 10.1097/MOG.0000000000000057
5. Vandeputte D, Falony G, Vieira-Silva S, Tito RY, Joossens M, Raes J. Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. *Gut.* (2016) 65:57–62. doi: 10.1136/gutjnl-2015-309618
6. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature.* (2010) 464:59. doi: 10.1038/nature08821
7. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. *Nature.* (2012) 486:222. doi: 10.1038/nature11053
8. Sonnenburg JL, Xu J, Leip DD, Chen C-H, Westover BP, Weatherford J, et al. Glycan foraging *in vivo* by an intestine-adapted bacterial symbiont. *Science.* (2005) 307:1955–9. doi: 10.1126/science.1109051
9. Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermúdez-Humarán LG, Gratadoux J-J, et al. *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci USA.* (2008) 105:16731–6. doi: 10.1073/pnas.0804812105
10. Belkaid Y, Hand TW. Role of the microbiota in immunity and inflammation. *Cell.* (2014) 157:121–41. doi: 10.1016/j.cell.2014.03.011
11. Chey WD, Kurlander J, Eswaran S. Irritable bowel syndrome: a clinical review. *JAMA.* (2015) 313:949–58. doi: 10.1001/jama.2015.0954
12. Sharon G, Sampson TR, Geschwind DH, Mazmanian SK. The central nervous system and the gut microbiome. *Cell.* (2016) 167:915–32. doi: 10.1016/j.cell.2016.10.027
13. Halfvarson J, Brislawn CJ, Lamendella R, Vázquez-Baeza Y, Walters WA, Bramer LM, et al. Dynamics of the human gut microbiome in inflammatory bowel disease. *Nat Microbiol.* (2017) 2:17004. doi: 10.1038/nmicrobiol.2017.4
14. Liu R, Hong J, Xu X, Feng Q, Zhang D, Gu Y, et al. Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. *Nat Med.* (2017) 23:859. doi: 10.1038/nm.4358
15. Wirbel J, Pyl PT, Kartal E, Zych K, Kashani A, Milanese A, et al. Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. *Nat Med.* (2019) 25:679. doi: 10.1038/s41591-019-0406-6
16. Suttle CA. Marine viruses—major players in the global ecosystem. *Nat Rev Microbiol.* (2007) 5:801. doi: 10.1038/nrmicro1750
17. Breitbart M. Marine viruses: truth or dare. *Ann Rev Mar Sci.* (2011) 4:425–48. doi: 10.1146/annurev-marine-120709-142805
18. Von Wintersdorff CJ, Penders J, Van Niekerk JM, Mills ND, Majumder S, Van Alphen LB, et al. Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. *Front Microbiol.* (2016) 7:173. doi: 10.3389/fmicb.2016.00173
19. Hurwitz BL, Sullivan MB. The Pacific Ocean Virome (POV): a marine viral metagenomic dataset and associated protein clusters for quantitative viral ecology. *PLoS ONE.* (2013) 8:e57355. doi: 10.1371/journal.pone.0057355
20. Hurwitz BL, Brum JR, Sullivan MB. Depth-stratified functional and taxonomic niche specialization in the 'core'and 'flexible'Pacific Ocean Virome. *ISME J.* (2015) 9:472. doi: 10.1038/ismej.2014.143
21. Lauro FM, Mcdougald D, Thomas T, Williams TJ, Egan S, Rice S, et al. The genomic basis of trophic strategy in marine bacteria. *Proc Natl Acad Sci USA.* (2009) 106:15527–33. doi: 10.1073/pnas.0903507106
22. Sullivan MB, Lindell D, Lee JA, Thompson LR, Bielawski JP, Chisholm SW. Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. *PLoS Biol.* (2006) 4:e234. doi: 10.1371/journal.pbio.0040234
23. Muniesa M, Hammerl JA, Hertwig S, Appel B, Brüssow H. Shiga toxin-producing *Escherichia coli* O104: H4: a new challenge for microbiology. *Appl Environ Microbiol.* (2012) 78:4065–73. doi: 10.1128/AEM.00217-12
24. Gregory AC, Zablocki O, Howell A, Bolduc B, Sullivan MB. The human gut virome database. *BioRxiv [Preprint].* (2019). doi: 10.1101/655910
25. Norman JM, Handley SA, Baldridge MT, Droit L, Liu CY, Keller BC, et al. Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell.* (2015) 160:447–60. doi: 10.1016/j.cell.2015.01.002
26. Manrique P, Bolduc B, Walk ST, Van Der Oost J, De Vos WM, Young MJ. Healthy human gut phageome. *Proc Natl Acad Sci USA.* (2016) 113:10400–5. doi: 10.1073/pnas.1601060113
27. Zuo T, Lu XJ, Zhang Y, Cheung CP, Lam S, Zhang F, et al. Gut mucosal virome alterations in ulcerative colitis. *Gut.* (2019) 68:1–11. doi: 10.1136/gutjnl-2018-318131
28. Hoyles L, Mccartney AL, Neve H, Gibson GR, Sanderson JD, Heller KJ, et al. Characterization of virus-like particles associated with the human faecal and caecal Characterization of virus-like particles associated with the human faecal and caecal microbiota. *Res Microbiol.* (2014) 165:803–12. doi: 10.1016/j.resmic.2014.10.006
29. Shkoporov AN, Ryan FJ, Draper LA, Forde A, Stockdale SR, Daly KM, et al. Reproducible protocols for metagenomic analysis of human faecal phageomes. *Microbiome.* (2018) 6:68. doi: 10.1186/s40168-018-0446-z
30. Breitbart M, Hewson I, Felts B, Mahaffy JM, Nulton J, Salamon P, et al. Metagenomic analyses of an uncultured viral community from human feces. *J Bacteriol.* (2003) 185:6220–3. doi: 10.1128/JB.185.20.6220-6223.2003
31. Aggarwala V, Liang G, Bushman FD. Viral communities of the human gut: metagenomic analysis of composition and dynamics. *Mobile DNA.* (2017) 8:12. doi: 10.1186/s13100-017-0095-y
32. Twort FW. An investigation on the nature of ultra-microscopic viruses. *Lancet.* (1915) 186:1241–3. doi: 10.1016/S0140-6736(01)20383-3
33. Forster SC, Kumar N, Anonye BO, Almeida A, Viciani E, Stares MD, et al. A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nat Biotechnol.* (2019) 37:186. doi: 10.1038/s41587-018-0009-7
34. Moreno-Gallego JL, Chou SP, Di Rienzi SC, Goodrich JK, Spector TD, Bell JT, et al. Virome diversity correlates with intestinal microbiome diversity in adult monozygotic twins. *Cell Host Microbe.* (2019) 25:261–72.e265. doi: 10.1016/j.chom.2019.01.019

35. Clooney AG, Sutton TDS, Shkoporov AN, Holohan RK, Daly KM, O'Regan O, et al. Whole-virome analysis sheds light on viral dark matter in inflammatory bowel disease. *Cell Host Microbe.* (2019). doi: 10.1016/j.chom.2019.10.009. [Epub ahead of print].
36. Roux S, Solonenko NE, Dang VT, Poulos BT, Schwenck SM, Goldsmith DB, et al. Towards quantitative viromics for both double-stranded and single-stranded DNA viruses. *PeerJ.* (2016) 4:e2777. doi: 10.7717/peerj.2777
37. Hesse U, Van Heusden P, Kirby BM, Olonade I, Van Zyl LJ, Trindade M. Virome assembly and annotation: a surprise in the Namib desert. *Front Microbiol.* 8:13. doi: 10.3389/fmicb.2017.00013
38. Sutton TD, Clooney AG, Ryan FJ, Ross RP, Hill C. Choice of assembly software has a critical impact on virome characterisation. *Microbiome.* (2019) 7:12. doi: 10.1186/s40168-019-0626-5
39. Weinbauer MG. Ecology of prokaryotic viruses. *FEMS Microbiol Rev.* (2004) 28:127–81. doi: 10.1016/j.femsre.2003.08.001
40. Ravin NV. Replication and maintenance of linear phage-plasmid N15. *Microbiol Spectr.* (2015) 3:71–82. doi: 10.1128/microbiolspec.PLAS-0032-2014
41. Cenens W, Makumi A, Govers SK, Lavigne R, Aertsen A. Viral transmission dynamics at single-cell resolution reveal transiently immune subpopulations caused by a carrier state association. *PLoS Genet.* (2016) 11:e1005770. doi: 10.1371/journal.pgen.1005770
42. Smeal SW, Schmitt MA, Pereira RR, Prasad A, Fisk JD. Simulation of the M13 life cycle I: assembly of a genetically-structured deterministic chemical kinetic simulation. *Virology.* (2017) 500:259–74. doi: 10.1016/j.virol.2016.08.017
43. Labrie SJ, Samson JE, Moineau S. Bacteriophage resistance mechanisms. *Nat Rev Microbiol.* (2010) 8:317. doi: 10.1038/nrmicro2315
44. Rostøl JT, Marraffini L. (Ph) ighting phages: how bacteria resist their parasites. *Cell Host Microbe.* (2019) 25:184–94. doi: 10.1016/j.chom.2019.01.009
45. Clement J-M, Lepouce E, Marchal C, Hofnung M. Genetic study of a membrane protein: DNA sequence alterations due to 17 lamB point mutations affecting adsorption of phage lambda. *EMBO J.* (1983) 2:77–80. doi: 10.1002/j.1460-2075.1983.tb01384.x
46. Chung IY, Jang HJ, Bae HW, Cho YH. A phage protein that inhibits the bacterial ATPase required for type IV pilus assembly. *Proc Natl Acad Sci USA.* (2014) 111:11503–8. doi: 10.1073/pnas.1403537111
47. Zago M, Orrù L, Rossetti L, Lamontanara A, Fornasari ME, Bonvini B, et al. Survey on the phage resistance mechanisms displayed by a dairy *Lactobacillus helveticus* strain. *Food Microbiol.* (2017) 66:110–6. doi: 10.1016/j.fm.2017.04.014
48. Scholl D, Adhya S, Merril C. *Escherichia coli* K1's capsule is a barrier to bacteriophage T7. *Appl Environ Microbiol.* (2005) 71:4872–4. doi: 10.1128/AEM.71.8.4872-4874.2005
49. Schwechheimer C, Kuehn MJ. Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions. *Nat Rev. Microbiol.* (2015) 13:605. doi: 10.1038/nrmicro3525
50. Cumby N, Reimer K, Mengin-Lecreulx D, Davidson AR, Maxwell KL. The phage tail tape measure protein, an inner membrane protein and a periplasmic chaperone play connected roles in the genome injection process of *E. coli* phage HK97. *Mol Microbiol.* (2015) 96:437–47. doi: 10.1111/mmi.12918
51. Tock MR, Dryden DT. The biology of restriction and anti-restriction. *Curr Opin Microbiol.* (2005) 8:466–72. doi: 10.1016/j.mib.2005.06.003
52. Kronheim S, Daniel-Ivad M, Duan Z, Hwang S, Wong AI, Mantel I, et al. A chemical defence against phage infection. *Nature.* (2018) 564:283–6. doi: 10.1038/s41586-018-0767-x
53. Ram G, Chen J, Kumar K, Ross HF, Ubeda C, Damle PK, et al. Staphylococcal pathogenicity island interference with helper phage reproduction is a paradigm of molecular parasitism. *Proc Natl Acad Sci USA.* (2012) 109:16300–5. doi: 10.1073/pnas.1204615109
54. Dy RL, Przybilski R, Semeijn K, Salmond GP, Fineran PC. A widespread bacteriophage abortive infection system functions through a Type IV toxin–antitoxin mechanism. *Nucleic Acids Res.* (2014) 42:4590–605. doi: 10.1093/nar/gkt1419
55. Van Houte S, Buckling A, Westra ER. Evolutionary ecology of prokaryotic immune mechanisms. *Microbiol Mol Biol Rev.* (2016) 80:745–63. doi: 10.1128/MMBR.00011-16
56. Hofer B, Ruge M, Dreiseikelmann B. The superinfection exclusion gene (sieA) of bacteriophage P22: identification and overexpression of the gene and localization of the gene product. *J Bacteriol.* (1995) 177:3080–6. doi: 10.1128/jb.177.11.3080-3086.1995
57. Samson JE, Magadán AH, Sabri M, Moineau S. Revenge of the phages: defeating bacterial defences. *Nat Rev Microbiol.* (2013) 11:675. doi: 10.1038/nrmicro3096
58. Leiman PG, Battisti AJ, Bowman VD, Stummeyer K, Mühlenhoff M, Gerardy-Schahn R, et al. The structures of bacteriophages K1E and K1-5 explain processive degradation of polysaccharide capsules and evolution of new host specificities. *J Mol Biol.* (2007) 371:836–49. doi: 10.1016/j.jmb.2007.05.083
59. Chatterjee S, Rothenberg E. Interaction of bacteriophage l with its *E. coli* receptor, LamB. *Viruses.* (2012) 4:3162–78. doi: 10.3390/v4113162
60. Minot S, Bryson A, Chehoud C, Wu GD, Lewis JD, Bushman FD. Rapid evolution of the human gut virome. *Proc Natl Acad Sci USA.* (2013) 110:12450–5. doi: 10.1073/pnas.1300833110
61. Warwick-Dugdale J, Solonenko N, Moore K, Chittick L, Gregory AC, Allen MJ, et al. Long-read viral metagenomics captures abundant and microdiverse viral populations and their niche-defining genomic islands. *PeerJ.* (2019) 7:e6800. doi: 10.7717/peerj.6800
62. Bondy-Denomy J, Pawluk A, Maxwell KL, Davidson AR. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature.* (2013) 493:429. doi: 10.1038/nature11723
63. Bair CL, Black LW. A type IV modification dependent restriction nuclease that targets glucosylated hydroxymethyl cytosine modified DNAs. *J Mol Biol.* (2007) 366:768–78. doi: 10.1016/j.jmb.2006.11.051
64. Seed KD, Lazinski DW, Calderwood SB, Camilli A. A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. *Nature.* (2013) 494:489. doi: 10.1038/nature11927
65. Scanlan PD. Bacteria–bacteriophage coevolution in the human gut: implications for microbial diversity and functionality. *Trends Microbiol.* (2017) 25:614–23. doi: 10.1016/j.tim.2017.02.012
66. Zhu A, Sunagawa S, Mende DR, Bork P. Inter-individual differences in the gene content of human gut bacterial species. *Genome Biol.* (2015) 16:82. doi: 10.1186/s13059-015-0646-9
67. Gandon S, Buckling A, Decaestecker E, Day T. Host–parasite coevolution and patterns of adaptation across time and space. *J Evol Biol.* (2008) 21:1861–6. doi: 10.1111/j.1420-9101.2008.01598.x
68. Hall AR, Scanlan PD, Morgan AD, Buckling A. Host-parasite coevolutionary arms races give way to fluctuating selection. *Ecol Lett.* (2011) 14:635–42. doi: 10.1111/j.1461-0248.2011.01624.x
69. De Sordi L, Lourenço M, Debarbieux L. "I will survive": a tale of bacteriophage-bacteria coevolution in the gut. *Gut Microbes.* (2019) 10:92–9. doi: 10.1080/19490976.2018.1474322
70. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* (2017) 27:824–34. doi: 10.1101/gr.213959.116
71. Porter NT, Hryckowian AJ, Merrill BD, Gardner JO, Singh S, Sonnenburg JL, et al. Multiple phase-variable mechanisms, including capsular polysaccharides, modify bacteriophage susceptibility in Bacteroides thetaiotaomicron. *BioRxiv [Preprint].* (2019). doi: 10.1101/521070
72. Turkington CJR, Morozov A, Clokie MRJ, Bayliss CD. Phage-resistant phase-variant sub-populations mediate herd immunity against bacteriophage invasion of bacterial meta-populations. *Front Microbiol.* 10. doi: 10.3389/fmicb.2019.01473
73. Denou E, Berger B, Barretto C, Panoff J-M, Arigoni F, Brüssow H. Gene expression of commensal *Lactobacillus johnsonii* strain NCC533 during *in vitro* growth and in the murine gut. *J Bacteriol.* (2007) 189:8109–19. doi: 10.1128/JB.00991-07
74. Maura D, Galtier M, Le Bouguénec C, Debarbieux L. Virulent bacteriophages can target O104: H4 enteroaggregative *Escherichia coli* in the mouse intestine. *Antimicrob Agents Chemother.* (2012) 56:6235–42. doi: 10.1128/AAC.00602-12
75. Galtier M, Sordi LD, Sivignon A, De Vallée A, Maura D, Neut C, et al. Bacteriophages targeting adherent invasive Escherichia coli strains as a promising new treatment for Crohn's disease. *J Crohn Colitis.* (2017) 11:840–7. doi: 10.1093/ecco-jcc/jjw224

76. Mirzaei MK, Maurice CF. Ménage à trois in the human gut: interactions between host, bacteria and phages. *Nat Rev Microbiol.* (2017) 15:397. doi: 10.1038/nrmicro.2017.30
77. Thingstad TF. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnol Oceanogr.* (2000) 45:1320–8. doi: 10.4319/lo.2000.45.6.1320
78. Avrani S, Schwartz DA, Lindell D. Virus-host swinging party in the oceans: incorporating biological complexity into paradigms of antagonistic coexistence. *Mobile Genet Elem.* (2012) 2:88–95. doi: 10.4161/mge.20031
79. Shkoporov AN, Hill C. Bacteriophages of the human gut: the "known unknown" of the microbiome. *Cell Host Microbe.* (2019) 25:195–209. doi: 10.1016/j.chom.2019.01.017
80. Breitbart M, Haynes M, Kelley S, Angly F, Edwards RA, Felts B, et al. Viral diversity and dynamics in an infant gut. *Res Microbiol.* (2008) 159:367–73. doi: 10.1016/j.resmic.2008.04.006
81. Lim ES, Zhou Y, Zhao G, Bauer IK, Droit L, Ndao IM, et al. Early life dynamics of the human gut virome and bacterial microbiome in infants. *Nat Med.* (2015) 21:1228. doi: 10.1038/nm.3950
82. Reyes A, Haynes M, Hanson N, Angly FE, Heath AC, Rohwer F, et al. Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature.* (2010) 466:334. doi: 10.1038/nature09199
83. Silpe JE, Bassler BL. A host-produced quorum-sensing autoinducer controls a phage lysis-lysogeny decision. *Cell.* (2019) 176:268–80.e213. doi: 10.1016/j.cell.2018.10.059
84. Knowles B, Silveira CB, Bailey BA, Barott K, Cantu VA, Cobián-Güemes AG, et al. Lytic to temperate switching of viral communities. *Nature.* (2016) 531:466. doi: 10.1038/nature17193
85. Silveira CB, Rohwer FL. Piggyback-the-Winner in host-associated microbial communities. *NPJ Biofilms Microbio.* (2016) 2:16010. doi: 10.1038/npjbiofilms.2016.10
86. Barr JJ, Auro R, Furlan M, Whiteson KL, Erb ML, Pogliano J, et al. Bacteriophage adhering to mucus provide a non–host-derived immunity. *Proc Natl Acad Sci USA.* (2013) 110:10771–6. doi: 10.1073/pnas.1305923110
87. Barr JJ, Auro R, Sam-Soon N, Kassegne S, Peters G, Bonilla N, et al. Subdiffusive motion of bacteriophage in mucosal surfaces increases the frequency of bacterial encounters. *Proc Natl Acad Sci USA.* (2015) 112:13675–80. doi: 10.1073/pnas.1508355112
88. Zhao G, Droit L, Gilbert MH, Schiro FR, Didier PJ, Si X, et al. Virome biogeography in the lower gastrointestinal tract of rhesus macaques with chronic diarrhea. *Virology.* (2019) 527:77–88. doi: 10.1016/j.virol.2018.10.001
89. Waller AS, Yamada T, Kristensen DM, Kultima JR, Sunagawa S, Koonin EV, et al. Classification and quantification of bacteriophage taxa in human gut metagenomes. *ISME J.* (2014) 8:1391. doi: 10.1038/ismej.2014.30
90. Ma Y, You X, Mai G, Tokuyasu T, Liu C. A human gut phage catalog correlates the gut phageome with type 2 diabetes. *Microbiome.* (2018) 6:24. doi: 10.1186/s40168-018-0410-y
91. Mccann A, Ryan FJ, Stockdale SR, Dalmasso M, Blake T, Ryan CA, et al. Viromes of one year old infants reveal the impact of birth mode on microbiome diversity. *PeerJ.* (2018) 6:e4694. doi: 10.7717/peerj.4694
92. Kim KH, Bae JW. Amplification methods bias metagenomic libraries of uncultured single-stranded and double-stranded DNA viruses. *Appl Environ Microbiol.* (2011) 77:7663–8. doi: 10.1128/AEM.00289-11
93. Pérez-Brocal V, García-López R, Vázquez-Castellanos JF, Nos P, Beltrán B, Latorre A, et al. Study of the viral and microbial communities associated with Crohn's disease: a metagenomic approach. *Clin Transl Gastroenterol.* (2013) 4:e36. doi: 10.1038/ctg.2013.9
94. Wagner J, Maksimovic J, Farries G, Sim WH, Bishop RF, Cameron DJ, et al. Bacteriophages in gut samples from pediatric Crohn's disease patients: metagenomic analysis using 454 pyrosequencing. *Inflam Bowel Dis.* (2013) 19:1598–608. doi: 10.1097/MIB.0b013e318292477c
95. Lepage P, Colombet J, Marteau P, Sime-Ngando T, Doré J, Leclerc M. Dysbiosis in inflammatory bowel disease: a role for bacteriophages? *Gut.* (2008) 57:424–5. doi: 10.1136/gut.2007.134668
96. Reyes A, Blanton LV, Cao S, Zhao G, Manary M, Trehan I, et al. Gut DNA viromes of Malawian twins discordant for severe acute malnutrition. *Proc Natl Acad Sci USA.* (2015) 112:11941–6. doi: 10.1073/pnas.1514285112
97. Fernandes MA, Verstraete SG, Phan TG, Deng X, Stekol E, Lamere B, et al. Enteric virome and bacterial microbiota in children with ulcerative colitis and Crohn disease. *J Pediatr Gastroenterol Nutr.* (2019) 68:30–6. doi: 10.1097/MPG.0000000000002140
98. Hulo C, Masson P, Le Mercier P, Toussaint A. A structured annotation frame for the transposable phages: a new proposed family "Saltoviridae" within the Caudovirales. *Virology.* (2015) 477:155–63. doi: 10.1016/j.virol.2014.10.009
99. Bolduc B, Jang HB, Doulcier G, You Z-Q, Roux S, Sullivan MB. vConTACT: an iVirus tool to classify double-stranded DNA viruses that infect Archaea and Bacteria. *PeerJ.* (2017) 5:e3243. doi: 10.7717/peerj.3243
100. Barylski J, Enault F, Dutilh BE, Schuller MBP, Edwards RA, Gillis A, et al. Analysis of spounaviruses as a case study for the overdue reclassification of tailed bacteriophages. *BioRxiv [Preprint].* (2018). doi: 10.1101/220434
101. Iranzo J, Krupovic M, Koonin EV. The double-stranded DNA virosphere as a modular hierarchical network of gene sharing. *mBio.* (2016) 7:e00978–16. doi: 10.1128/mBio.00978-16
102. Jang HB, Bolduc B, Zablocki O, Kuhn JH, Roux S, Adriaenssens EM, et al. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nat Biotechnol.* (2019) 37:632. doi: 10.1038/s41587-019-0100-8
103. Dutilh BE, Schmieder R, Nulton J, Felts B, Salamon P, Edwards RA, et al. Reference-independent comparative metagenomics using cross-assembly: crAss. *Bioinformatics.* (2012) 28:3225–31. doi: 10.1093/bioinformatics/bts613
104. Dutilh BE, Cassman N, Mcnair K, Sanchez SE, Silva GG, Boling L, et al. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nat Commun.* (2014) 5:5498. doi: 10.1038/ncomms5498
105. Edwards RA, Vega AA, Norman HM, Ohaeri M, Levi K, Dinsdale EA, et al. Global phylogeography and ancient evolution of the widespread human gut virus crAssphage. *Nat Microbiol.* (2019) 4:1727–36. doi: 10.1038/s41564-019-0494-6
106. Shkoporov AN, Khokhlova EV, Fitzgerald CB, Stockdale SR, Draper LA, Ross RP, et al. ΦCrAss001 represents the most abundant bacteriophage family in the human gut and infects Bacteroides intestinalis. *Nat Commun.* (2018) 9:4781. doi: 10.1038/s41467-018-07225-7
107. Guerin E, Shkoporov A, Stockdale SR, Clooney AG, Ryan FJ, Sutton TDS, et al. Biology and taxonomy of crass-like bacteriophages, the most abundant virus in the human gut. *Cell Host Microbe.* (2018) 24:653–64.e656. doi: 10.1016/j.chom.2018.10.002
108. Yutin N, Makarova KS, Gussow AB, Krupovic M, Segall A, Edwards RA, et al. Discovery of an expansive bacteriophage family that includes the most abundant viruses from the human gut. *Nat Microbiol.* (2018) 3:38. doi: 10.1038/s41564-017-0053-y
109. Draper LA, Ryan FJ, Smith MK, Jalanka J, Mattila E, Arkkila PA, et al. Long-term colonisation with donor bacteriophages following successful faecal microbial transplantation. *Microbiome.* (2018) 6:220. doi: 10.1186/s40168-018-0598-x
110. Siringan P, Connerton PL, Cummings NJ, Connerton IF. Alternative bacteriophage life cycles: the carrier state of *Campylobacter jejuni*. *Open Biol.* (2014) 4:130200. doi: 10.1098/rsob.130200
111. Breitbart M, Bonnain C, Malki K, Sawaya NA. Phage puppet masters of the marine microbial realm. *Nat Microbiol.* (2018) 3:754–66. doi: 10.1038/s41564-018-0166-y
112. De Sordi L, Khanna V, Debarbieux L. The gut microbiota facilitates drifts in the genetic diversity and infectivity of bacterial viruses. *Cell Host Microbe.* (2017) 22:801–8.e803. doi: 10.1016/j.chom.2017.10.010
113. Hurwitz BL, Hallam SJ, Sullivan MB. Metabolic reprogramming by viruses in the sunlit and dark ocean. *Genome Biol.* (2013) 14:R123. doi: 10.1186/gb-2013-14-11-r123
114. Waldor MK, Mekalanos JJ. Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science.* (1996) 272:1910–4. doi: 10.1126/science.272.5270.1910
115. Modi SR, Lee HH, Spina CS, Collins JJ. Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. *Nature.* (2013) 499:219. doi: 10.1038/nature12212

116. Enault F, Briet A, Bouteille L, Roux S, Sullivan MB, Petit M-A. Phages rarely encode antibiotic resistance genes: a cautionary tale for virome analyses. *ISME J.* (2017) 11:237–47. doi: 10.1038/ismej.2016.90
117. Roux S, Krupovic M, Debroas D, Forterre P, Enault F. Assessment of viral community functional potential from viral metagenomes may be hampered by contamination with cellular sequences. *Open Biol.* (2013) 3:130160. doi: 10.1098/rsob.130160
118. Sutton TDS, Clooney AG, Hill C. Giant oversights in the human gut virome. *Gut*. (2019). doi: 10.1136/gutjnl-2019-319067
119. Brüssow H, Canchaya C, Hardt W-D. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev.* (2004) 68:560–602. doi: 10.1128/MMBR.68.3.560-602.2004
120. Obeng N, Pratama AA, Van Elsas JD. The significance of mutualistic phages for bacterial ecology and evolution. *Trends Microbiol.* (2016) 24:440–9. doi: 10.1016/j.tim.2015.12.009
121. Wright A. Mechanism of conversion of the Salmonella O antigen by bacteriophage ε34. *J Bacteriol.* (1971) 105:927–36.
122. Verma N, Brandt J, Verma D, Lindberg A. Molecular characterization of the O-acetyl transferase gene of converting bacteriophage SF6 that adds group antigen 6 to Shigella flexneri. *Mol Microbiol.* (1991) 5:71–5. doi: 10.1111/j.1365-2958.1991.tb01827.x
123. Davies MR, Broadbent SE, Harris SR, Thomson NR, Van Der Woude MW. Horizontally acquired glycosyltransferase operons drive salmonellae lipopolysaccharide diversity. *PLoS Genet.* (2013) 9:e1003568. doi: 10.1371/journal.pgen.1003568
124. Diard M, Bakkeren E, Cornuault JK, Moor K, Hausmann A, Sellin ME, et al. Inflammation boosts bacteriophage transfer between *Salmonella* spp. *Science.* (2017) 355:1211–5. doi: 10.1126/science.aaf8451
125. Chen J, Quiles-Puchalt N, Chiang YN, Bacigalupe R, Fillol-Salom A, Chee MSJ, et al. Genome hypermobility by lateral transduction. *Science.* (2018) 362:207–12. doi: 10.1126/science.aat5867
126. Castro-Mejía JL, Muhammed MK, Kot W, Neve H, Franz CM, Hansen LH, et al. Optimizing protocols for extraction of bacteriophages prior to metagenomic analyses of phage communities in the human gut. *Microbiome.* (2015) 3:64. doi: 10.1186/s40168-015-0131-4
127. Lang AS, Beatty JT. Importance of widespread gene transfer agent genes in α-proteobacteria. *Trends Microbiol.* (2007) 15:54–62. doi: 10.1016/j.tim.2006.12.001
128. Shkoporov AN, Clooney AG, Sutton TDS, Ryan FJ, Daly KM, Nolan JA, et al. The human gut virome is highly diverse, stable, and individual specific. *Cell Host Microbe.* (2019) 26:527–41.e525. doi: 10.1016/j.chom.2019.09.009
129. Minot S, Wu GD, Lewis JD, Bushman FD. Conservation of gene cassettes among diverse viruses of the human gut. *PLoS ONE.* (2012) 7:e42342. doi: 10.1371/journal.pone.0042342
130. Monaco CL, Gootenberg DB, Zhao G, Handley SA, Ghebremichael MS, Lim ES, et al. Altered virome and bacterial microbiome in human immunodeficiency virus-associated acquired immunodeficiency syndrome. *Cell Host Microbe*. (2016) 19:311–22. doi: 10.1016/j.chom.2016.02.011
131. Zhao G, Vatanen T, Droit L, Park A, Kostic AD, Poon TW, et al. Intestinal virome changes precede autoimmunity in type I diabetes-susceptible children. *Proc Natl Acad Sci USA*. (2017) 114:E6166–75. doi: 10.1073/pnas.1706359114
132. Hannigan GD, Duhaime MB, Ruffin MT, Koumpouras CC, Schloss PD. Diagnostic potential and interactive dynamics of the colorectal cancer virome. *mBio.* (2018) 9:e02248–18. doi: 10.1128/mBio.02248-18
133. Dinh DM, Volpe GE, Duffalo C, Bhalchandra S, Tai AK, Kane AV, et al. Intestinal microbiota, microbial translocation, and systemic inflammation in chronic HIV infection. *J Infect Dis.* (2014) 211:19–27. doi: 10.1093/infdis/jiu409
134. Gogokhia L, Buhrke K, Bell R, Hoffman B, Brown DG, Hanke-Gogokhia C, et al. Expansion of bacteriophages is linked to aggravated intestinal inflammation and colitis. *Cell Host Microbe.* (2019) 25:285–99.e288. doi: 10.1016/j.chom.2019.01.008
135. Van Belleghem JD, Clement F, Merabishvili M, Lavigne R, Vaneechoutte M. Pro- and anti-inflammatory responses of peripheral blood mononuclear cells induced by *Staphylococcus aureus* and *Pseudomonas aeruginosa* phages. *Sci Rep*. (2017) 7:8004. doi: 10.1038/s41598-017-08336-9
136. Ivanenkov VV, Menon AG. Peptide-mediated transcytosis of phage display vectors in MDCK cells. *Biochem Biophys Res Commun.* (2000) 276:251–7. doi: 10.1006/bbrc.2000.3358
137. Nguyen S, Baker K, Padman BS, Patwa R, Dunstan RA, Weston TA, et al. Bacteriophage transcytosis provides a mechanism to cross epithelial cell layers. *mBio.* (2017) 8:e01874–17. doi: 10.1128/mBio.01874-17
138. Górski A, Wazna E, Dabrowska BW, Dabrowska K, Switała-Jelen K, Miedzybrodzki R. Bacteriophage translocation. *FEMS Immunol Med Microbiol.* (2006) 46:313–9. doi: 10.1111/j.1574-695X.2006.00044.x
139. Górski A, Dabrowska K, Miedzybrodzki R, Weber-Dabrowska B, Łusiak-Szelachowska M, Jonczyk-Matysiak E, et al. Phages and immunomodulation. *Fut Microbiol.* (2017) 12:905–14. doi: 10.2217/fmb-2017-0049
140. Łusiak-Szelachowska M, Weber-Dabrowska B, Jonczyk-Matysiak E, Wojciechowska R, Górski A. Bacteriophages in the gastrointestinal tract and their implications. *Gut Pathogens.* (2017) 9:44. doi: 10.1186/s13099-017-0196-7
141. Górski A, Jonczyk-Matysiak E, Miedzybrodzki R, Weber-Dabrowska B, Łusiak-Szelachowska M, Baginska N, et al. Phage therapy: beyond antibacterial action. *Front Med.* 5:146. doi: 10.3389/fmed.2018.00146
142. Sweere JM, Van Belleghem JD, Ishak H, Bach MS, Popescu M, Sunkari V, et al. Bacteriophage trigger antiviral immunity and prevent clearance of bacterial infection. *Science*. (2019) 363:eaat9691. doi: 10.1126/science.aat9691
143. Jalanka J, Mattila E, Jouhten H, Hartman J, De Vos WM, Arkkila P, et al. Long-term effects on luminal and mucosal microbiota and commonly acquired taxa in faecal microbiota transplantation for recurrent Clostridium difficile infection. *BMC Med.* (2016) 14:155. doi: 10.1186/s12916-016-0698-z
144. Cammarota G, Ianiro G, Gasbarrini A. Fecal microbiota transplantation for the treatment of *Clostridium difficile* infection: a systematic review. *J Clin Gastroenterol.* (2014) 48:693–702. doi: 10.1097/MCG.0000000000000046
145. Ott SJ, Waetzig GH, Rehman A, Moltzau-Anderson J, Bharti R, Grasis JA, et al. Efficacy of sterile fecal filtrate transfer for treating patients with *Clostridium difficile* infection. *Gastroenterology*. (2017) 152:799–811.e797. doi: 10.1053/j.gastro.2016.11.010
146. Moineau S, Lévesque C. Control of bacteriophages in industrial fermentations. *Bacteriophages.* (2005) 285–96. doi: 10.1201/9780203491751.ch10
147. Wilson BC, Vatanen T, Cutfield WS, O'sullivan JM. The super-donor phenomenon in fecal microbiota transplantation. *Front Cell Infect Microbiol*. (2019) 9:2. doi: 10.3389/fcimb.2019.00002
148. Conceição-Neto N, Deboutte W, Dierckx T, Machiels K, Wang J, Yinda KC, et al. Low eukaryotic viral richness is associated with faecal microbiota transplantation success in patients with UC. *Gut*. (2018) 67:1558–9. doi: 10.1136/gutjnl-2017-315281
149. Vaughn BP, Vatanen T, Allegretti JR, Bai A, Xavier RJ, Korzenik J, et al. Increased intestinal microbial diversity following fecal microbiota transplant for active Crohn's disease. *Inflam Bowel Dis.* (2016) 22:2182–90. doi: 10.1097/MIB.0000000000000893
150. Ng SC, Shi HY, Hamidi N, Underwood FE, Tang W, Benchimol EI, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *Lancet.* (2017) 390:2769–78. doi: 10.1016/S0140-6736(17)32448-0
151. Pérez-Brocal V, García-López R, Nos P, Beltrán B, Moret I, Moya A. Metagenomic analysis of Crohn's disease patients identifies changes in the virome and microbiome related to disease status and therapy, and detects potential interactions and biomarkers. *Inflamm Bowel Dis.* (2015) 21:2515–32. doi: 10.1097/MIB.0000000000000549
152. Casjens SR, Hendrix RW. Bacteriophage lambda: early pioneer and still relevant. *Virology*. (2015) 479–480:310–30. doi: 10.1016/j.virol.2015.02.010
153. Górski A, Miedzybrodzki R, Borysowski J, Dabrowska K, Wierzbicki P, Ohams M, et al. Chapter 2–phage as a modulator of immune responses: practical implications for phage therapy. In: Łobocka M, Szybalski W, editors. *Advances in Virus Research.* San Diego, CA: Elsevier Academic Press Inc. (2012), p. 41–71.
154. King AM, Lefkowitz E, Adams MJ, Carstens EB. *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*. San Diego, CA: Elsevier Academic Press Inc. (2011).
155. Roux S, Emerson JB, Eloe-Fadrosh EA, Sullivan MB. Benchmarking viromics: an *in silico* evaluation of metagenome-enabled estimates

of viral community composition and diversity. *PeerJ.* (2017) 5:e3817. doi: 10.7717/peerj.3817
156. Hannigan GD, Meisel JS, Tyldsley AS, Zheng Q, Hodkinson BP, Sanmiguel AJ, et al. The human skin double-stranded DNA virome: topographical and temporal diversity, genetic enrichment, and dynamic associations with the host microbiome. *MBio.* (2015) 6:e01578–15. doi: 10.1128/mBio.01578-15
157. Yilmaz S, Allgaier M, Hugenholtz P. Multiple displacement amplification compromises quantitative analysis of metagenomes. *Nat Methods.* (2010) 7:943. doi: 10.1038/nmeth1210-943
158. Probst AJ, Weinmaier T, Desantis TZ, Santo Domingo JW, Ashbolt N. New perspectives on microbial community distortion after whole-genome amplification. *PLoS ONE.* (2015) 10:e0124158. doi: 10.1371/journal.pone.0124158
159. Zhang K, Martiny AC, Reppas NB, Barry KW, Malek J, Chisholm SW, et al. Sequencing genomes from single cells by polymerase cloning. *Nat Biotechnol.* (2006) 24:680–6. doi: 10.1038/nbt1214
160. Woyke T, Xie G, Copeland A, González JM, Han C, Kiss H, et al. Assembling the marine metagenome, one cell at a time. *PLoS ONE.* (2009) 4:e5299. doi: 10.1371/journal.pone.0005299
161. García-López R, Vázquez-Castellanos JF, Moya A. Fragmentation and coverage variation in viral metagenome assemblies, and their effect in diversity calculations. *Front Bioeng Biotechnol.* 3:141. doi: 10.3389/fbioe.2015.00141
162. Vollmers J, Wiegand S, Kaster AK. Comparing and evaluating metagenome assembly tools from a microbiologist's perspective-Not only size matters! *PLoS ONE.* (2017) 12:e0169662. doi: 10.1371/journal.pone.0169662
163. Paez-Espino D, Eloe-Fadrosh EA, Pavlopoulos GA, Thomas AD, Huntemann M, Mikhailova N, et al. Uncovering Earth's virome. *Nature.* (2016) 536:425. doi: 10.1038/nature19094
164. Simmonds P, Adams MJBenko M, Breitbart M, Brister JR, Carstens EB, et al. Virus taxonomy in the age of metagenomics. *Nat Rev Microbiol.* (2017) 15:161. doi: 10.1038/nrmicro.2016.177
165. Aiewsakun P, Adriaenssens EM, Lavigne R, Kropinski AM, Simmonds P. Evaluation of the genomic diversity of viruses infecting bacteria, archaea and eukaryotes using a common bioinformatic platform: steps towards a unified taxonomy. *J Gen Virol.* (2018) 99:1331–43. doi: 10.1099/jgv.0.001110
166. Eloe-Fadrosh EA. Towards a genome-based virus taxonomy. *Nat Microbiol.* (2019) 4:1249–50. doi: 10.1038/s41564-019-0511-9
167. Simmonds P, Aiewsakun P. Virus classification–where do you draw the line? *Arch Virol.* (2018) 163:2037–46. doi: 10.1007/s00705-018-3938-z
168. Friman VP, Hiltunen T, Jalasvuori M, Lindstedt C, Laanto E, Örmälä AM, et al. High temperature and bacteriophages can indirectly select for bacterial pathogenicity in environmental reservoirs. *PLoS ONE.* (2011) 6:e17651. doi: 10.1371/journal.pone.0017651
169. Scanlan PD, Buckling A, Hall AR. Experimental evolution and bacterial resistance: (co)evolutionary costs and trade-offs as opportunities in phage therapy research. *Bacteriophage.* (2015) 5:e1050153. doi: 10.1080/21597081.2015.1050153
170. Brüssow H. Hurdles for phage therapy to become a reality—an editorial comment. *Viruses.* (2019) 11:E557. doi: 10.3390/v11060557
171. Wright A, Hawkins CH, Änggård EE, Harper DR. A controlled clinical trial of a therapeutic bacteriophage preparation in chronic otitis due to antibiotic-resistant Pseudomonas aeruginosa; a preliminary report of efficacy. *Clin Otolaryngol.* (2009) 34:349–57. doi: 10.1111/j.1749-4486.2009.01973.x
172. Chan BK, Turner PE, Kim S, Mojibian HR, Elefteriades JA, Narayan D. Phage treatment of an aortic graft infected with Pseudomonas aeruginosa. *Evol Med Public Health.* (2018) 2018:60–6. doi: 10.1093/emph/eoy005
173. Lavergne S, Hamilton T, Biswas B, Kumaraswamy M, Schooley RT, Wooten D. Phage therapy for a multidrug-resistant *Acinetobacter baumannii* craniectomy site infection. *Open Forum Infect Dis.* (2008) 5:ofy064. doi: 10.1093/ofid/ofy064
174. Garrett L. Seven circles of antimicrobial hell. *Lancet.* (2019) 393:865–7. doi: 10.1016/S0140-6736(19)30397-6

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Sutton and Hill. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.