ANNUAL REVIEWS

Annual Review of Microbiology

Factors Affecting Variation of the Human Gut Phageome

Ciara A. Tobin,[1,2] Colin Hill,[1,2] and Andrey N. Shkoporov[1,2,3]

[1]APC Microbiome Ireland, Cork, Ireland; email: 116439372@umail.ucc.ie, c.hill@ucc.ie, andrey.shkoporov@ucc.ie

[2]School of Microbiology, University College Cork, Cork, Ireland

[3]Department of Medicine, University College Cork, Cork, Ireland

Annu. Rev. Microbiol. 2023. 77:363–79

First published as a Review in Advance on June 12, 2023

The *Annual Review of Microbiology* is online at micro.annualreviews.org

https://doi.org/10.1146/annurev-micro-032421-105754

Keywords

bacteriophage, gut microbiome, virome, health, disease

Abstract

The gut microbiome is a dense and metabolically active consortium of microorganisms and viruses located in the lower gastrointestinal tract of the human body. Bacteria and their viruses (phages) are the most abundant members of the gut microbiome. Investigating their biology and the interplay between the two is important if we are to understand their roles in human health and disease. In this review, we summarize recent advances in resolving the taxonomic structure and ecological functions of the complex community of phages in the human gut—the gut phageome. We discuss how age, diet, and geography can all have a significant impact on phageome composition. We note that alterations to the gut phageome have been observed in several diseases such as inflammatory bowel disease, irritable bowel syndrome, and colorectal cancer, and we evaluate whether these phageome changes can directly or indirectly contribute to disease etiology and pathogenesis. We also highlight how lack of standardization in studying the gut phageome has contributed to variation in reported results.

Contents

INTRODUCTION

The human gut microbiome is a complex and dynamic environment, estimated to be composed of $\sim 10^{14}$ microbial cells from all domains of cellular life (79). These inhabitants of the gastrointestinal tract play a role in various important metabolic processes and are vital for normal human gut physiology. Bacteria have received the greatest attention, likely because they have the largest biomass and account for ~93% of DNA in metagenomic studies (82). Numerous studies have highlighted their importance in both health and disease, while the viral, fungal, and protozoan components are often neglected. The viral component of the gut includes both eukaryotic viruses and bacteriophages (phages), the latter being the most abundant members of the gut virome. Phages, together with their bacterial hosts, are the most abundant components of the gut microbiome. In recent years, viral metagenomics has revolutionized research of complex viral communities and has enabled us to catalog hundreds of thousands of previously unknown phages of the gut (7, 12, 67). The expansion of cataloged phages has also resulted in numerous taxonomic changes ratified by the International Committee on Taxonomy of Viruses (ICTV) (47). One of the most recent changes has been the abolishment of the previous Siphoviridae, Myoviridae, and Podoviridae families as well as the Caudovirales order, with all members assigned to the class *Caudoviricetes* (96, 97). This is in recognition of the inaccuracy of using morphology-based taxonomy and moves instead toward genome-based classification.

Phages persist in the gut microbiome at $\sim 10^9$–10^{10} viral-like particles (VLPs) per gram of feces. They are estimated to exist at an ~1:1 ratio with their bacterial hosts, although it has been suggested that viruses are outnumbered by bacteria in the gut, with ratios as low as 0.1:1 (82). As such, they can potentially exert considerable influence over bacteria through a variety of mechanisms such as lysis, lysogenic conversion, chronic infection, and phage transduction (horizontal gene transfer) (10). This can lead to both direct and indirect impacts on the human host in both health and disease. It has been noted that the complex community of gut phages, referred to as the gut phageome, is highly variable between individuals while being temporally stable within individuals (81). In this review, we analyze the state of the art of phageome research, in particular focusing on phageome variation and its implications in both health and disease states (**Figure 1**).

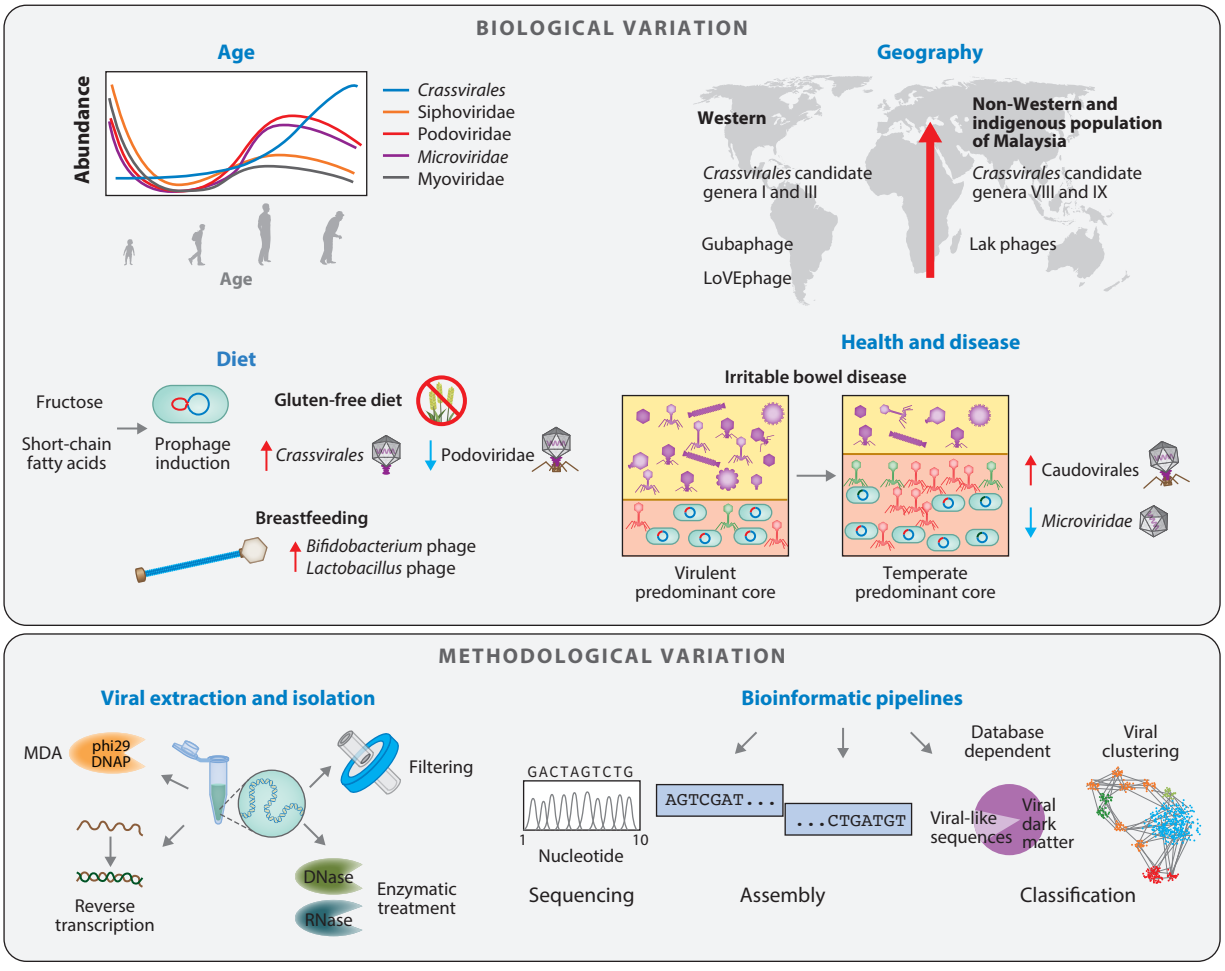

Figure 1

Biological and methodological factors contributing to phageome variation in the human gut. Biological factors include age, geography, diet, and various health and disease states. Methodological factors include viral extraction and isolation protocols and bioinformatic pipelines used. Abbreviations: DNAP, DNA polymerase; MDA, multiple displacement amplification. Figure adapted from images created with BioRender.com.

PHAGE-BACTERIAL HOST DYNAMICS

Historically, both the lytic and lysogenic phage life cycles were expected to operate in the gut microbiome. Both involve a phage particle identifying and adsorbing to a receptor on the surface of the bacterial cell and injecting its genetic material into the susceptible host. In the lytic life cycle, carried out by virulent phage, the phage takes over the host cell machinery to replicate its own genetic material. Once mature virions are assembled, the bacterial cell is lysed to release new progeny. In lysogeny, carried out by temperate phage, the genetic material of the phage also has the possibility of integrating into the host chromosome as a prophage. This prophage can be replicated for generations in tandem with bacterial reproduction until it is induced into the lytic life cycle, which is usually due to some environmental stressor but can also occur spontaneously (**Figure 2**). These are the main phage infection strategies observed to occur in the gut (21). Pseudolysogeny,

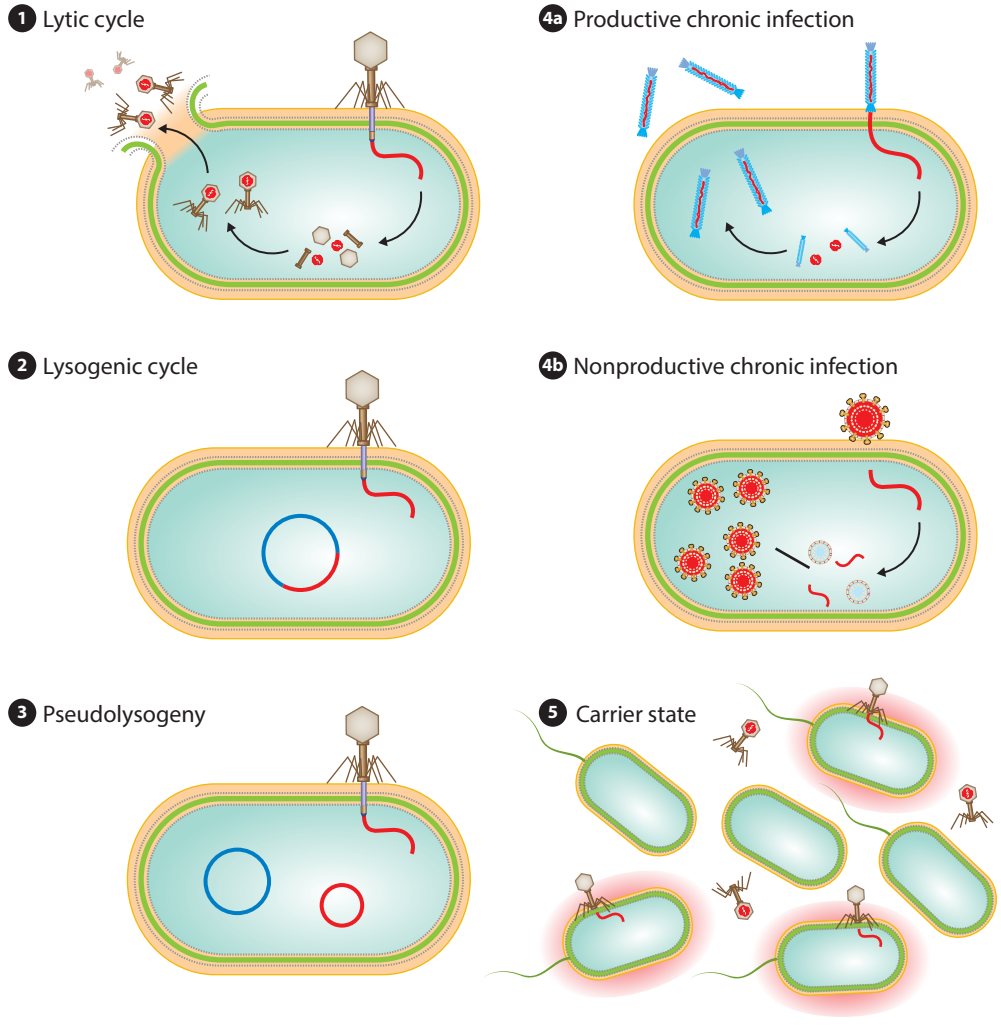

Figure 2

The principal life cycles that have been described for bacteriophages: (❶) lytic cycle, (❷) lysogenic cycle, (❸) pseudolysogeny, (❹) productive and nonproductive chronic infection, and (❺) carrier state. In the carrier state depicted here, a subpopulation of bacteria revert to resistance through loss of flagella, which can act as a phage receptor, and switching expression to a different capsular polysaccharide. Figure adapted from images created with BioRender.com.

carrier state, and chronic infection are the most common alternative phage infection strategies to be described. However, they are poorly characterized, with the terms pseudolysogeny and carrier state being used interchangeably in literature (58). As such, there have been efforts to define each more clearly (38, 58).

Pseudolysogeny has been described as a form of delayed phage reproduction in which the phage genetic material remains dormant in the bacterial cell. It has been noted to occur in phages T4 (55) and P22 (14), which infect *Escherichia coli* and *Salmonella enterica* serovar Typhimurium. Chronic infection describes the ability of certain phage infections to continuously produce virions in the bacterial cell. In the productive form, these are excreted from the cell without the requirement to lyse it. In the nonproductive form, assembled phages accumulate in the cytoplasm of the bacterial

cell but are not released. Productive chronic infection is most commonly described as occurring in filamentous phages, while nonproductive chronic infection has been reported in RNA phages (58).

While these four life cycles occur at the single-cell level, it has been suggested that the carrier state should be viewed at the population level, with both phage and their bacterial hosts persisting in equilibrium. A subpopulation of resistant bacteria serves to continuously maintain this balance. The carrier state has been described for *Campylobacter jejuni* and two of its phages (87), while the persistence of phages of the recently established order *Crassvirales* (88, 97) and their bacterial hosts has also been suggested to involve a population-level carrier state (83).

The temporal stability and persistence observed in the gut would suggest that temperate bacteria dominate the gut microbiome. Contrary to this, a large proportion of the gut phageome is composed of the virulent *Microviridae*, as well as viruses previously classified as Podoviridae, and *Crassvirales* (29, 31, 81). It is possible that alternative phage–bacterial host interactions play a larger role in the gut microbiome than currently anticipated, with benign interactions maintaining population stability. This shifts the view of phages as primarily predators of bacteria to participants in a more mutualistic relationship, in which both partners can benefit (85). In addition to this, spatial heterogeneity (56), host resistance, and phage counter-resistance mechanisms (32) and various abiotic and biotic factors along the gastrointestinal tract can all influence the maintenance of bacteria-phage populations and interactions.

BIOLOGICAL FACTORS CONTRIBUTING TO PHAGEOME VARIATION

Age

While the concept of a healthy core phageome has been examined in adults, it varies quite extensively over an entire life span. Phage populations in infants are quite unstable and dynamic, reflecting the early establishment of the gut bacteriome (the bacterial members of the gut microbiome). Epifluorescent microscopy analysis of meconium and stool samples from the first days of life reveal very few or no VLPs (11, 53). In the first month of life, the infant virome undergoes a dramatic expansion, reaching up to 10^9 VLPs per gram of stool, which is consistent with counts in healthy adults (11, 39). Viral richness has in general been observed to increase over time from birth and strongly correlates with bacterial richness and diversity (6, 53). However, Lim and colleagues (54) noted a steady decrease in both phage richness and diversity in fecal samples taken from infants and children over a period of two years. During the first year of life, the gut bacteriome has been described to go through three developmental phases, described as gut microbiota maturation stages. This has been characterized by shifts from *Escherichia* to *Bifidobacterium* and then to *Bacteroidetes* (5). It has been observed that over 60% of early infant phages associate with one of these gut microbiota maturation stages (6). Clostridia have also been noted to continuously increase in relative abundance over the first two years of life (54). Phages previously classified as Siphoviridae and Myoviridae, along with eukaryotic viruses, tend to dominate in infants but steadily decline in absolute abundance into childhood (31, 92). Corresponding with the abundance of Siphoviridae phages, the vast majority of phages identified in infant stool samples have been predicted to have a temperate lifestyle (6). The relative abundance of these temperate phages also reduces over time, with mothers having a significantly lower level compared to their infants. This reduction, together with strong temporal correlations of phages and their bacterial hosts, has supported the hypothesis that the initial viral colonizers of the infant gut are prophages that excise from their bacterial hosts (6, 53). Although lytic *Crassvirales* are the most abundant phages in the adult gut, they are only present at a low abundance in the infant gut, often remaining absent in reads from fecal samples (6, 31, 53).

While several studies have focused on the dynamics of the infant gut virome, less attention has been given to the viromes of healthy children. In an analysis of gut viral richness across different age groups, using data sets from several studies, Gregory and colleagues (31) found that viral richness is lowest during childhood, before significantly increasing into adulthood. However, in a study on the gut virome of Danish children and adults, this significant increase in viral richness could not be recapitulated (98). Limitations in both papers, including uneven sample sizes (31) and a lack of data from young adults (98), may have contributed to the discrepancies in these results. During the transition from childhood to adulthood, the previously dominant phages classified as Siphoviridae and Myoviridae are outcompeted by Podoviridae, *Microviridae*, and *Crassvirales* (31). Bacterial composition remains relatively stable during this period, although slight shifts in the relative abundances of the phyla *Bacteroidetes* and *Actinobacteria*, with an increase in the former and a decrease in the latter, have been observed (71). A longitudinal study of ten healthy adults over a period of one year saw a similar persistence of phage families, although phages previously classified as Podoviridae were less abundant (81). In general, age appears to have a positive correlation with virome diversity, increasing from young adults to elderly adults in a Japanese cohort (69). However, in Western populations, this positive correlation is lost: Phage populations begin to decrease again as people reach old age. This is with the exception of *Crassvirales*, whose abundance continues to increase across the life span and appears to be at its highest in the elderly population (31). In an analysis of the gut viromes of over 4,000 Japanese individuals, age was shown to have the strongest association with the virome among 232 host and environmental factors analyzed. One hundred and seventy-six viral operational taxonomic units were significantly associated with age. Among these, phages of *Ruminococcus*, *Faecalibacterium*, and *Clostridium* were enriched in elderly people, while a viral cluster (VC) of *Bifidobacterium*-related phages were associated with young adults and steadily decreased with age (69).

Diet

It is well established that diet has a significant impact on the bacteriome. High-fat and Western diets, which have been associated with obesity, cardiovascular disease, and type 2 diabetes, have been shown to negatively impact bacterial gut diversity and function. Other diets associated with health, such as gluten-free, vegan, and vegetarian diets, have also resulted in differing bacterial compositions (95). Given the influence of diet on the abundance and diversity of phage hosts, it is a fair assumption that diet will exert similar influences on the phageome.

In infants, differences in phageome composition have been observed based on whether the infant was fed formula or breastfed. Specifically, temperate phages of *Bifidobacterium* and *Lactobacillus* are increased in breastfed infants, with a corresponding enrichment of their hosts (53). In an analysis of the viromes of 930 healthy Chinese adults, diet was found to have a significant effect on virome variation, with 22 dietary components contributing to significant variation. One hundred and thirty-seven correlations between dietary components and viral species could be made (106). Similarly, diet contributed to virome variation in Japanese individuals, with products such as milk, coffee, and fruit significantly associated (69). The interplay of ethnicity, diet, and virome variation was also highlighted in the former study (106). *Bacteroides* phages were noted to be enriched in the viromes of the Zang ethnic group as compared to the five other ethnic groups analyzed in this study. Butter milk tea and barley were predominant features in the diets of ethnic Zang, with barley significantly correlating with *Bacteroides* phage and butter milk tea being the primary contributor to bacterial variation (106). A limited number of studies involving deliberate dietary intervention have suggested that diet can have some effect on the composition of the virome. Implementation of a week-long gluten free-diet in 11 healthy adults resulted in shifts in phage family abundance, decreasing phages previously classified as Podoviridae and increasing *Crassvirales*. At

the viral genomic contig level, the number of shared contigs among individuals increased during a gluten-free diet (29). A study of high- and low-fat dietary interventions also gave rise to the idea that a shared diet can increase similarity in gut virome composition (63). A high-fat diet in mice significantly depletes phages previously classified as Siphoviridae (77). Concomitant with this is a significant reduction in integrase genes. Given that a high proportion of phages classified as Siphoviridae have a temperate lifestyle, these results suggest that a high-fat diet may shift the core virome toward expansion of lytic phage. However, this contrasts with findings by Kim & Bae (44), who observed an expansion of temperate phages previously classified as Caudovirales in mice on a high-fat, high-sucrose diet. Both fructose and short-chain fatty acids have been shown to activate prophage induction from *Lactobacillus reuteri* in the gut of mice (72). This induction appears to be partly due to the activation of stress responses in *L. reuteri*. In addition, many common dietary components were demonstrated to result in prophage induction from several gut bacteria (9), highlighting the potential influence that diet can exert on both the gut phageome and the microbiome as a whole.

Geography

Several recent studies have consistently noted differences in gut phageome composition based on geographical location (12, 65, 98, 106). Two of the most prevalent phage genomes identified in a Danish cohort, a *Crassvirales* phage and a newly identified LoVEphage, were also recognized to be prolific worldwide. However, their abundance differed by location (98). The *Crassvirales* phage was found to be more abundant in healthy people in Asia and Europe than in healthy people in the United States and Africa. The LoVEphage was more prevalent in healthy Europeans and Americans than in healthy Asians and Africans. Thirty-seven other prevalent phage genomes were also significantly associated with geographical location, being found most commonly in Europeans (98). Similarly, another recently uncovered, highly abundant phage clade, Gubaphage, is most prolific in Europeans and absent in samples originating from South America (12). The relative abundance of ten candidate genera of *Crassvirales* varies significantly depending on geographical location. The most prevalent in Western populations is candidate genus I, which contains prototypical crAssphage (p-crAssphage) (named after the cross-assembly software originally used to identify it), and candidate genus III (12, 33). In Africa and South America, candidate genera VIII and IX are much more abundant. In the largest analysis of globally distributed p-crAssphage sequences, p-crAssphage was found to cluster based on geographical location (26). The gut viromes differed significantly between individuals living in two different regions of China, Hong Kong and Yunnan (106). All alpha diversity indices were significantly lower in the Hong Kong population: They were depleted for several phages previously classified as Caudovirales while being enriched for *Microviridae*. Urbanization also appears to play a role in variation, with significant differences in virome structure between individuals inhabiting urban and rural regions of Yunnan. The length of time spent in an urban region was also positively associated with several phages, including phages of lactococci and lactobacilli (106).

In a study examining diarrheal disease in an Indian cohort, individuals dwelling in the urban region of Nagpur had a less diverse virome than those living in the rural area of Melghat (65). It was also found that geographical location exerted more of an influence over virome variation than diarrheal disease did. It was also found to have an impact on gut virome alterations in obese patients from China. Obesity appeared to have more of an effect on the gut virome of individuals living in Hong Kong than it did on the Kunming cohort. There was a significant reduction in gut viral richness and diversity in obese Hong Kong individuals, but this significance was lost in the individuals from Kunming (100).

With geography demonstrated to have a considerable impact on the gut virome, it would be prudent to characterize the composition of healthy gut viromes across the world. Currently, gut virome studies are biased toward Europe and the United States, while coverage of Asia is dominated by studies focused on China and Japan. Studies focusing on South America and Africa are lacking. In recent efforts to characterize the viromes of indigenous populations in Malaysia, researchers found that their virome composition was markedly different from other viromes worldwide (49). A focus on analyzing the gut viromes of individuals from Bangladesh resulted in the discovery of Lak phages, termed megaphages due to their extremely large genomes (>540 kb) (24). *Prevotella* species dominated in individuals in which these phages were identified, and these bacteria were identified as likely Lak phage hosts via CRISPR targeting. *Prevotella* is prevalent in indigenous and non-Western microbiomes, which is attributed to the consumption of a high-fiber, low-fat diet (35). This discovery highlights the importance of moving away from a Western-dominated view of the gut phageome.

PHAGEOME VARIABILITY IN DISEASE

Inflammatory Bowel Disease

Inflammatory bowel disease (IBD) is a term used to describe two chronic inflammatory conditions of the gut: ulcerative colitis (UC) and Crohn's disease. Although there is no known etiology, some causative factors have been linked, including genetics, diet, environmental exposures, and gut microbiome composition (46, 60, 103). A handful of studies have focused on changes to the gut virome in IBD patients, where alterations to the phageome have been observed, but the variations noted differ between studies. The most consistent observation is the increase in abundance of phages previously classified as Caudovirales in individuals with IBD, whether this be in relation to healthy controls (70, 99, 105) or in comparison across disease type (27), although no expansion of these phages has also been noted (52). An increase in the richness and diversity of phages previously classified as Caudovirales has also been reported (17, 70), while Zuo and colleagues (105) observed a significant reduction in the richness and diversity of phages previously classified as Caudovirales. The viromes of mucosa samples taken from the rectum were examined here, rather than fecal samples, highlighting how sample sites can also contribute to variation. Multiple studies have also reported an altered ratio of the abundance of phages previously classified as Caudovirales to that of *Microviridae*: The relative abundance of phages previously classified as Caudovirales is significantly increased in IBD patients (52, 70). A large quantity of phages previously classified as Caudovirales have been identified to be enriched and associated with IBD. Most commonly these have been phages of *Lactococcus*, *Lactobacillus*, *Clostridium*, *Enterococcus*, and *Escherichia* (40, 52, 70, 105). Imai and colleagues (40) observed a significant increase in abundance of *Crassvirales* in Crohn's disease patients. It should be noted that this pattern was observed in a cohort of only 19 patients. In contrast, an analysis of *Crassvirales* and their association with diseases found that 19 genus-level clusters of *Crassvirales* were significantly depleted in a cohort of >2,000 IBD patients (34). In addition, Clooney et al. (17) identified a depletion of a *Crassvirales* VC to be associated with a shift to IBD status.

Many studies use the viral-like sequences identified at the species level to determine virome-specific changes in disease. However, interpersonal gut virome variation can result in changes to the virome being masked at this level of analysis. In addition, often only those viral sequences that can be taxonomically assigned are used for analysis, despite this being as little as just 5% of the total in some studies (76). To address this, Clooney et al. (17) reanalyzed the data from Norman et al.'s study (70) and reduced taxonomic resolution by clustering viral sequences to family level (8). The majority of VCs found to be significantly increased in IBD patients were predicted to

contain temperate phages, while the highly abundant VCs in controls were largely lytic in nature (17). Similarly, Sinha et al. (86) observed that an inoculum of pooled VLPs from UC patients contained a significantly higher proportion of temperate phages than that of healthy controls. These data suggest that in IBD patients, the core virome of virulent phages associated with healthy individuals (81) shifts to being predominantly temperate. It is likely that stressors associated with the inflamed environment of an IBD gut result in prophage induction. Importantly, this shift in composition was not detectable when analysis was performed at species level.

In addition to phageome variation, Zuo et al. (105) reported an increase in *Chrysochromulina ericina virus* and *Mimivirus*, two giant eukaryotic viruses of the family *Mimiviridae*, in their UC cohort. However, the use of 0.22-μm filters, followed by lysozyme and chloroform treatment during VLP extraction, means that the likelihood of *Mimivirus*, with a capsid size of 0.5 μm, ending up in the extracted VLPs is extremely low (89). However, other studies that have used similar filter size and treatment steps have also reported the presence of *Mimiviridae* in the gut virome (77, 100). The presence of these viruses likely reflects incorrect taxonomic assignment (89) and highlights one of the challenges of characterizing the human gut virome. It is also possible that virophages, which are known to infect *Mimiviridae* and carry *Mimiviridae*-like elements, are present and contribute to incorrect taxonomic assignment (73). Any correlation made between *Mimiviridae* and IBD or other diseases should be interpreted with caution.

Recent work in mice has hinted at a more direct role of altered phages in IBD. The addition of *E. coli* phages to the drinking water of specific pathogen-free mice was shown to exacerbate the severity of induced colitis. Phage interaction with the immune system, mediated by IFN-γ induction, resulted in this increase in inflammation. These in vivo findings were further validated when VLPs isolated from patients with active UC disease were shown to induce IFN-γ production from $CD4^+$ T cells (30). Similarly, mice colonized by bacterial communities from UC patients had increased proinflammatory cytokines and exacerbated colitis, when they underwent gavage with pooled VLPs from UC patients (86). In addition, UC VLPs were found to have an effect on the UC bacterial community, including reducing the relative abundance of *Eubacterium limosum* and increasing the relative abundance of *Escherichia* and *Shigella* species. *E. limosum* has previously been shown to improve experimental colitis symptoms (42), while pathobiont strains of *E. coli* are frequently implemented in IBD development (64). Taken together, these findings suggest a tripartite role of bacteria, phages, and the immune system in IBD development.

Irritable Bowel Syndrome

Irritable bowel syndrome (IBS) is a common gastrointestinal disorder, estimated to affect up to 21% of the general population worldwide (16). Despite links to various host and environmental factors, there is no identified causative factor. Diagnosis is based predominantly on symptoms and by exclusion of other inflammatory diseases. Rome IV criteria divide it into four clinical subtypes based on the primary symptoms: diarrhea (IBS-D), constipation (IBS-C), a mix of diarrhea and constipation (IBS-M), or unclassified (IBS-U) (48).

Four recent studies examining the gut virome in IBS patients have consistently reported variations to the gut phageome (3, 19, 50, 62). In general, IBS patients have a less rich and diverse virome compared to controls (3, 19), although Mihindukulasuriya et al. (62) did not find any difference in alpha diversity metrics. Coughlan and colleagues (19) reported that gut viromes of IBS-D and IBS-M but not IBS-C patients remained significantly less diverse than those of controls. Ansari and colleagues' (3) data set also contained these subtypes, while Mihindukulasuriya et al. had a data set of IBS-C and IBS-D patients only. Differences in subtypes may account for the differing diversity metrics reported and highlight how variation, even within the subtypes of disease, can influence reported results.

Coughlan et al. (19) identified ten VCs that were differentially abundant between IBS patients and controls. Seven were of reduced abundance in IBS patients, including two VCs previously classified as Siphoviridae and one previously classified as Podoviridae, while one VC previously classified as Podoviridae and one previously classified as Siphoviridae were increased in the IBS cohort (19). At the viral species level, Mihindukulasuriya et al. (62) also identified variation in abundance of ten phages classified as Siphoviridae, Myoviridae, Podoviridae, and *Microviridae*. They noted that an increase in these phages could also be stratified by IBS subtype. Li et al. (50) reported that the enrichment of viruses in IBS patients correlated with an enrichment of bacteria and that the same pattern occurred for the depletion of viruses and bacteria, signaling a possible interlinked role for the two in IBS. Unlike the latter three studies, Ansari et al. (3) did not find any significant alterations to the phageome. Instead they noted an increase in abundance of several large eukaryotic viruses such as *Pandoravirus salinus* and *Herpesviridae* and a decreased abundance of several eukaryotic viruses that typically infect hosts other than humans, such as plants and amoebae.

The shift from a dominant virulent core to one that is predominantly temperate that is noted to occur in the gut phageome of patients with IBD (17) was not observed for patients with IBS (19). This emphasizes that despite IBD and IBS being somewhat similar symptomatically, there are likely different alterations and etiological factors at play. A couple of these studies were limited by a relatively small cohort of IBS patients, so that splitting by IBS subtype was not feasible. Given that there have been several indications that the IBS subtype contributes to variation, future studies should consider inclusion and analysis of all subtypes. Investigation into the link between the gut virome and IBS is still in its infancy, and future studies will hopefully elucidate whether changes are a causative factor of IBS or occur simply as a consequence of the disease.

Colorectal Cancer

Colorectal cancer (CRC) is the third-most common cancer diagnosis. Development of this cancer is complex and believed to be multifactorial, with risks including genetic and environmental factors and lifestyle choices (23, 59). Alterations to the gut microbiome are increasingly recognized as an important environmental risk factor. Significant deviation to overall virome structure has been observed in CRC patients (28, 80). In general, viral diversity and richness appear to increase in CRC patients (66, 80, 107), although in some studies alpha diversity metrics were not able to detect any differences (28, 36). Despite this, every gut virome study of CRC patients has identified virome signatures associated with CRC, the majority of which are a result of changes in phage composition. In general, enrichment of phages classified as Siphoviridae, Myoviridae, Podoviridae, *Drexlerviridae*, and *Inoviridae* was identified to be driving this signature (36, 66, 107). Nakatsu and colleagues (66) reported that phages from the genera *Inovirus* and *Tunalikevirus* were the greatest discriminators between people with CRC and healthy controls. Gao et al. (28) identified an increase in abundance in phages infecting *Escherichia* and *Salmonella* in CRC patients, while p-crAssphage and *Enterobacteria* phage were depleted in these patients when compared to those in healthy controls. Enrichment of a *Parabacteroides* phage was also associated with a mutated KRAS gene, a frequently mutated gene in CRC that is linked with poorer prognosis and recovery (28, 61). *Herelleviridae*, a relatively newly identified family of phages that infect *Bacillota* (formerly Firmicutes), are also noted to be diminished in CRC patients (4, 107).

Whether these shifts in phage composition are directly involved in CRC etiology is unknown. Enrichment of *Inoviridae* in CRC is intriguing, as they have been linked with biofilm development (78), a factor that has been associated with progression of CRC (22). Changes in a number of key bacterial species have been linked to CRC, some of which have proposed mechanisms for CRC progression. Enrichment of *Fusobacterium nucleatum* has been most commonly identified in several

studies (37). This bacterium has the ability to promote tumorigenesis through the invasion of damaged gut epithelial cells and interacting with Toll-like receptors in the tumor cell (45). Expansion of enterotoxigenic *Bacteroides fragilis*, *E. coli* carrying the polyketide synthase gene, *Parvimonas micra*, and bacteria of the genera *Porphyromonas* and *Peptostreptococcus* has also been commonly reported (15). These bacteria are potential hosts to several of the phages previously noted to be enriched in CRC (28, 66). In addition, Shen et al. (80) identified phages of *F. nucleatum* and *P. micra* as potential biomarkers of CRC. This suggests a possible interlink of bacteria and phage in the development of CRC. On the other hand, a low correlation in the abundance of the identified influential bacteria and phages has been noted (36). This has led to one hypothesis that alterations to the phageome indirectly impact CRC development. This would be through altering the previously stable bacterial community, providing the opportunity for the expansion of bacterial drivers of CRC (36). In a study analyzing gut virome changes during induced tumor formation in mice, phages of the genera *Brunovirus* and *Hpunavirus* were positively associated with tumor growth, hinting at a possible more direct involvement in CRC development (51). While these alterations in the phageome highlight the possible use of these phages as CRC biomarkers, further studies are required to elucidate the extent of the role of the phageome in CRC development and progression.

METHODOLOGICAL VARIATION

It is evident that there is enormous variation in the human gut phageome of individuals. However, many metagenomic studies focus their attention on changes that occur in the identifiable fraction of the gut virome. The remaining unidentified component, termed "viral dark matter" (2), is likely to have a large impact on many facets of the gut microbiome and as demonstrated by Clooney et al. (17) can change the outcome of studies when it is included in analysis. Unlike 16S rRNA or the *chaperonin-60* (*cpn60*) gene in bacteria, viruses have no universal phylogenetic marker, leaving researchers to rely on sequence homology searches of reference databases. The primary reference database used, NCBI RefSeq, has only 11,598 viral reference genomes (as of September 2022). Rapid mutation rates and mosaicism also lead to many highly divergent sequences unlikely to be present in databases, further hindering identification. To overcome this, many studies incorporate de novo assembly of reads, which investigators can then attempt to annotate and classify through a combination of de novo and reference-based steps. The choice of assembler has also been shown to impact virome identification and classification (90).

Perhaps the best example of the requirement to move away from using only reference-based identification was the discovery of *Crassvirales*. They remained a component of the large unidentified mass of viral reads until 2014, when the use of cross-assembly software (crAss) led to the identification of p-crAssphage (25). Since then, a plethora of *Crassvirales* have been identified from diverse environments (26, 101, 102). Highly prolific and widespread, they can account for over 86% of viral reads from some individuals and are a core component of the virome in healthy adults (101). Recent efforts based on de novo assembly from globally distributed metagenomes have led to the establishment of several human gut virome databases, expanding our knowledge on both the abundance and diversity of gut virome genomes (12, 31, 67, 93). The establishment of both these and future databases will be beneficial for alignment-based analysis. The assembly of a catalog of gut virome databases has also led to the identification of several new phages. The highly prevalent LoVEphage, named as such because it has lots of viral elements, was discovered through the assembly of a Danish gut viral catalog (98). The assembly of the Gut Phage Database, the largest amalgamation of gut phage genomes so far, led to identification of a clade of phage named Gubaphage. With many similar features to *Crassvirales*, they were also found to be highly ubiquitous, appearing in samples from all continents except for South America (12).

Although suggestions of standard protocols have been put forward (18, 84), there is currently no standardized approach to studying the gut phageome. This leads to inevitable variation in results, as researchers adopt different protocols. It also hinders accurate comparisons across studies. There are two main methods through which viral sequences are isolated: (*a*) viral extraction from stool and enrichment and (*b*) a bulk metagenomics approach. While bulk metagenomics has the advantages of a simpler isolation protocol and capturing temperate phages, rare phages are less likely to be recovered (75). Recent gut virome studies have seen an increase in the use of bulk metagenomics (3, 20, 65, 94), and comparison between the two methods has demonstrated that a greater number of viral sequences are detected from bulk metagenomics (31). Continuous improvements to identification of viral sequences with bulk metagenomics (41) further increase the appeal of this method. However, comparison of the two methods also shows only ~10% of identified viral sequences shared between enriched and bulk viromes (31), highlighting how the two methods select for very different portions of the virome. This also suggests that an integrated approach of both VLP enrichment and bulk metagenomics should be used to capture greater viral diversity and representation. In the enrichment and isolation of VLPs, there is often a trade-off between maximizing the number of VLPs obtained and minimizing the amount of bacterial contamination. Filters with a pore size of 0.22–0.45 μm are commonly used to reduce bacterial contamination but at the loss of large and enveloped viruses. Chloroform can provide further purification from bacterial cells but also results in the loss of some chloroform-sensitive viruses. The use of RNase to remove free nucleic acids could potentially have a negative impact on RNA viruses, leading to further bias toward enrichment of DNA viruses (1). In addition to this, many studies do not include a reverse transcription step to include RNA viruses and focus on the DNA portion alone. The recent expansion of metagenomic sequences predicted to be RNA phages further suggests that the number of gut RNA phages is being underestimated (68). Extraction and enrichment of nucleic acids often result in a low yield and require an amplification step. Multiple displacement amplification is most often used in studies (31), despite being well documented to introduce bias via amplification of ssDNA viruses. This is due to the preference of the enzyme used, polymerase φ29, to propagate ssDNA (43). It is likely to be a contributing factor as to why some studies report a high relative abundance of *Microviridae*.

The choice of sequencing platform is also important, as sequencing depth, chemistry, and read length can all further complicate the analysis of virome variation. This can be seen particularly in some of the earlier IBD studies that were restricted by the sequencing depth of 454 pyrosequencing and therefore found limited changes to the gut virome (74, 99). Due to much lower DNA concentrations being required, most gut virome studies use short-read sequencing platforms. Short-read sequencing also has the advantage of allowing deep sequencing but can hinder downstream analysis, as short reads can often lead to fragmented assemblies (91). The use of long-read sequencing could overcome these assembly inaccuracies, but the requirement of a high starting concentration of DNA limits its application in gut virome analysis. Recently, Lee et al. (49) used an optimized PEG-precipitation protocol to achieve on average 20 ng/μL of DNA from fecal samples of indigenous Malaysians. This high starting concentration could be used for long-read sequencing, which despite a lower sequencing depth was shown to improve the breadth of coverage compared to short-read sequencing.

CONCLUSIONS

Many other biological and health and disease states have also been associated with alterations to the gut phageome. Sex significantly correlates with the gut phageome, with males enriched in phages that infect *Megamonas* and *Prevotella* while females have increased abundances of

Faecalibacterium and *Ruminococcus* phages (69). *Crassvirales* have been found to be significantly depleted in patients with two autoimmune diseases, rheumatoid arthritis and systemic lupus erythematosus (94). Along with other phageome alterations, *Crassvirales* were significantly reduced in patients infected with SARS-CoV-2 (13, 57, 104). Phage VCs were significantly altered in individuals with metabolic syndrome (20). At the same time, an analysis of 232 host and environmental factors contributed only 0.6% of variance to the gut virome in a cohort of healthy Japanese individuals. This was much lower than the 5% of variance of the bacteriome that could be attributed to these factors (69), suggesting that there are still many unknown components contributing to virome variation between individuals.

To understand alterations to the gut virome in disease states, it is vital that we have a clear understanding of virome composition in healthy individuals. Currently our view of a healthy phageome is predominantly based on Western adults. Data from children and elderly people are lacking, as well as data from people in large portions of Africa, Asia, and South America. It is important that future work focus on incorporating data from these under-sampled groups in order to form as complete an understanding as possible of the gut phageome. This understanding is also necessary because of the increasing use of phage therapy and other medical applications such as fecal microbiota/viral transplantation.

Standardized protocols are required to enable cross study comparisons. At the same time, this should not dissuade researchers from pursuing novel techniques, as these are likely to further advance the field of gut virome research. Despite great progress in the last couple of years, gut phageome research is still in its infancy, particularly compared to the efforts made in deciphering the gut bacteriome. This is a dynamic field of research, and as methods improve and studies adapt to improve gaps in gut phageome knowledge, it is likely that current views will change.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

1. Adriaenssens EM, Farkas K, Harrison C, Jones DL, Allison HE, McCarthy AJ. 2018. Viromic analysis of wastewater input to a river catchment reveals a diverse assemblage of RNA viruses. *mSystems* 3:e00025-18
2. Aggarwala V, Liang G, Bushman FD. 2017. Viral communities of the human gut: metagenomic analysis of composition and dynamics. *Mob. DNA* 8:12
3. Ansari MH, Ebrahimi M, Fattahi MR, Gardner MG, Safarpour AR, et al. 2020. Viral metagenomic analysis of fecal samples reveals an enteric virome signature in irritable bowel syndrome. *BMC Microbiol.* 20:123
4. Barylski J, Kropinski AM, Alikhan NF, Adriaenssens EM, ICTV Rep. Consort. 2020. ICTV virus taxonomy profile: *Herelleviridae*. *J. Gen. Virol.* 101:362–63
5. Beller L, Deboutte W, Falony G, Vieira-Silva S, Tito RY, et al. 2021. Successional stages in infant gut microbiota maturation. *mBio* 12:e0185721
6. Beller L, Deboutte W, Vieira-Silva S, Falony G, Tito RY, et al. 2022. The virota and its transkingdom interactions in the healthy infant gut. *PNAS* 119:e2114619119
7. Benler S, Yutin N, Antipov D, Rayko M, Shmakov S, et al. 2021. Thousands of previously unknown phages discovered in whole-community human gut metagenomes. *Microbiome* 9:78
8. Bin Jang H, Bolduc B, Zablocki O, Kuhn JH, Roux S, et al. 2019. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nat. Biotechnol.* 37:632–39
9. Boling L, Cuevas DA, Grasis JA, Kang HS, Knowles B, et al. 2020. Dietary prophage inducers and antimicrobials: toward landscaping the human gut microbiome. *Gut Microbes* 11:721–34

10. Borodovich T, Shkoporov AN, Ross RP, Hill C. 2022. Phage-mediated horizontal gene transfer and its implications for the human gut microbiome. *Gastroenterol. Rep.* 10:goac012
11. Breitbart M, Haynes M, Kelley S, Angly F, Edwards RA, et al. 2008. Viral diversity and dynamics in an infant gut. *Res. Microbiol.* 159:367–73
12. Camarillo-Guerrero LF, Almeida A, Rangel-Pineros G, Finn RD, Lawley TD. 2021. Massive expansion of human gut bacteriophage diversity. *Cell* 184:1098–109.e9
13. Cao J, Wang C, Zhang Y, Lei G, Xu K, et al. 2021. Integrated gut virome and bacteriome dynamics in COVID-19 patients. *Gut Microbes* 13:1887722
14. Cenens W, Makumi A, Mebrhatu MT, Lavigne R, Aertsen A. 2013. Phage-host interactions during pseudolysogeny: lessons from the Pid/dgo interaction. *Bacteriophage* 3:e25029
15. Cheng Y, Ling Z, Li L. 2020. The intestinal microbiota and colorectal cancer. *Front. Immunol.* 11:615056
16. Chey WD, Kurlander J, Eswaran S. 2015. Irritable bowel syndrome: a clinical review. *JAMA* 313:949–58
17. Clooney AG, Sutton TDS, Shkoporov AN, Holohan RK, Daly KM, et al. 2019. Whole-virome analysis sheds light on viral dark matter in inflammatory bowel disease. *Cell Host Microbe* 26:764–78.e5
18. Conceição-Neto N, Zeller M, Lefrère H, De Bruyn P, Beller L, et al. 2015. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Sci. Rep.* 5:16532
19. Coughlan S, Das A, O'Herlihy E, Shanahan F, O'Toole PW, Jeffery IB. 2021. The gut virome in Irritable Bowel Syndrome differs from that of controls. *Gut Microbes* 13:1
20. de Jonge PA, Wortelboer K, Scheithauer TPM, van den Born BH, Zwinderman AH, et al. 2022. Gut virome profiling identifies a widespread bacteriophage family associated with metabolic syndrome. *Nat. Commun.* 13:3594
21. De Sordi L, Lourenço M, Debarbieux L. 2019. The battle within: interactions of bacteriophages and bacteria in the gastrointestinal tract. *Cell Host Microbe* 25:210–18
22. Dejea CM, Wick EC, Hechenbleikner EM, White JR, Mark Welch JL, et al. 2014. Microbiota organization is a distinct feature of proximal colorectal cancers. *PNAS* 111:18321–26
23. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. 2019. Colorectal cancer. *Lancet* 394:1467–80
24. Devoto AE, Santini JM, Olm MR, Anantharaman K, Munk P, et al. 2019. Megaphages infect *Prevotella* and variants are widespread in gut microbiomes. *Nat. Microbiol.* 4:693–700
25. Dutilh BE, Cassman N, McNair K, Sanchez SE, Silva GG, et al. 2014. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nat. Commun.* 5:4498
26. Edwards RA, Vega AA, Norman HM, Ohaeri M, Levi K, et al. 2019. Global phylogeography and ancient evolution of the widespread human gut virus crAssphage. *Nat. Microbiol.* 4:1727–36
27. Fernandes MA, Verstraete SG, Phan TG, Deng X, Stekol E, et al. 2019. Enteric virome and bacterial microbiota in children with ulcerative colitis and Crohn disease. *J. Pediatr. Gastroenterol. Nutr.* 68:30–36
28. Gao R, Zhu Y, Kong C, Xia K, Li H, et al. 2021. Alterations, interactions, and diagnostic potential of gut bacteria and viruses in colorectal cancer. *Front. Cell Infect. Microbiol.* 11:657867
29. Garmaeva S, Gulyaeva A, Sinha T, Shkoporov AN, Clooney AG, et al. 2021. Stability of the human gut virome and effect of gluten-free diet. *Cell Rep.* 35:109132
30. Gogokhia L, Buhrke K, Bell R, Hoffman B, Brown DG, et al. 2019. Expansion of bacteriophages is linked to aggravated intestinal inflammation and colitis. *Cell Host Microbe* 25:285–99.e8
31. Gregory AC, Zablocki O, Zayed AA, Howell A, Bolduc B, Sullivan MB. 2020. The Gut Virome Database reveals age-dependent patterns of virome diversity in the human gut. *Cell Host Microbe* 28:724–40.e8
32. Guerin E, Hill C. 2020. Shining light on human gut bacteriophages. *Front. Cell Infect. Microbiol.* 10:481
33. Guerin E, Shkoporov A, Stockdale SR, Clooney AG, Ryan FJ, et al. 2018. Biology and taxonomy of crAss-like bacteriophages, the most abundant virus in the human gut. *Cell Host Microbe* 24:653–64.e6
34. Gulyaeva A, Garmaeva S, Ruigrok R, Wang D, Riksen NP, et al. 2022. Discovery, diversity, and functional associations of crAss-like phages in human gut metagenomes from four Dutch cohorts. *Cell Rep.* 38:110204
35. Gupta VK, Paul S, Dutta C. 2017. Geography, ethnicity or subsistence-specific variations in human microbiome composition and diversity. *Front. Microbiol.* 8:1162
36. Hannigan GD, Duhaime MB, Ruffin MT 4th, Koumpouras CC, Schloss PD. 2018. Diagnostic potential and interactive dynamics of the colorectal cancer virome. *mBio* 9(6):e02248-18

37. Hashemi Goradel N, Heidarzadeh S, Jahangiri S, Farhood B, Mortezaee K, et al. 2019. *Fusobacterium nucleatum* and colorectal cancer: a mechanistic overview. *J. Cell Physiol.* 234:2337–44
38. Hobbs Z, Abedon ST. 2016. Diversity of phage infection types and associated terminology: the problem with 'Lytic or lysogenic'. *FEMS Microbiol. Lett.* 363(7):fnw047
39. Hoyles L, McCartney AL, Neve H, Gibson GR, Sanderson JD, et al. 2014. Characterization of virus-like particles associated with the human faecal and caecal microbiota. *Res. Microbiol.* 165(10):803–12
40. Imai T, Inoue R, Nishida A, Yokota Y, Morishima S, et al. 2022. Features of the gut prokaryotic virome of Japanese patients with Crohn's disease. *J. Gastroenterol.* 57:559–70
41. Johansen J, Plichta DR, Nissen JN, Jespersen ML, Shah SA, et al. 2022. Genome binning of viral entities from bulk metagenomics data. *Nat. Commun.* 13:965
42. Kanauchi O, Fukuda M, Matsumoto Y, Ishii S, Ozawa T, et al. 2006. *Eubacterium limosum* ameliorates experimental colitis and metabolite of microbe attenuates colonic inflammatory action with increase of mucosal integrity. *World J. Gastroenterol.* 12(7):1071–77
43. Kim KH, Bae JW. 2011. Amplification methods bias metagenomic libraries of uncultured single-stranded and double-stranded DNA viruses. *Appl. Environ. Microbiol.* 77:7663–68
44. Kim MS, Bae JW. 2016. Spatial disturbances in altered mucosal and luminal gut viromes of diet-induced obese mice. *Environ. Microbiol.* 18:1498–510
45. Kostic AD, Chun E, Robertson L, Glickman JN, Gallini CA, et al. 2013. *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe* 14:207–15
46. Krela-Kaźmierczak I, Zakerska-Banaszak O, Skrzypczak-Zielińska M, Łykowska-Szuber L, Szymczak-Tomczak A, et al. 2022. Where do we stand in the behavioral pathogenesis of inflammatory bowel disease? The Western dietary pattern and microbiota—a narrative review. *Nutrients* 14(12):2520
47. Krupovic M, Turner D, Morozova V, Dyall-Smith M, Oksanen HM, et al. 2021. Bacterial Viruses Subcommittee and Archaeal Viruses Subcommittee of the ICTV: update of taxonomy changes in 2021. *Arch. Virol.* 166:3239–44
48. Lacy BE, Patel NK. 2017. Rome criteria and a diagnostic approach to irritable bowel syndrome. *J. Clin. Med.* 6(11):99
49. Lee CZ, Zoqratt M, Phipps ME, Barr JJ, Lal SK, et al. 2022. The gut virome in two indigenous populations from Malaysia. *Sci. Rep.* 12:1824
50. Li M, Wang C, Guo Q, Xu C, Xie Z, et al. 2022. More positive or more negative? Metagenomic analysis reveals roles of virome in human disease-related gut microbiome. *Front. Cell Infect. Microbiol.* 12:846063
51. Li Y, Zhang F, Zheng H, Kalasabail S, Hicks C, et al. 2022. Fecal DNA virome is associated with the development of colorectal neoplasia in a murine model of colorectal cancer. *Pathogens* 11(4):457
52. Liang G, Conrad MA, Kelsen JR, Kessler LR, Breton J, et al. 2020. Dynamics of the stool virome in very early-onset inflammatory bowel disease. *J. Crohn's Colitis* 14:1600–10
53. Liang G, Zhao C, Zhang H, Mattei L, Sherrill-Mix S, et al. 2020. The stepwise assembly of the neonatal virome is modulated by breastfeeding. *Nature* 581:470–74
54. Lim ES, Zhou Y, Zhao G, Bauer IK, Droit L, et al. 2015. Early life dynamics of the human gut virome and bacterial microbiome in infants. *Nat. Med.* 21:1228–34
55. Los M, Wegrzyn G, Neubauer P. 2003. A role for bacteriophage T4 *rI* gene function in the control of phage development during pseudolysogeny and in slowly growing host cells. *Res. Microbiol.* 154:547–52
56. Lourenço M, Chaffringeon L, Lamy-Besnier Q, Pédron T, Campagne P, et al. 2020. The spatial heterogeneity of the gut limits predation and fosters coexistence of bacteria and bacteriophages. *Cell Host Microbe* 28:390–401.e5
57. Lu ZH, Zhou HW, Wu WK, Fu T, Yan M, et al. 2021. Alterations in the composition of intestinal DNA virome in patients with COVID-19. *Front. Cell Infect. Microbiol.* 11:790422
58. Mäntynen S, Laanto E, Oksanen HM, Poranen MM, Díaz-Muñoz SL. 2021. Black box of phage-bacterium interactions: exploring alternative phage infection strategies. *Open. Biol.* 11:210188
59. Mármol I, Sánchez-de-Diego C, Pradilla Dieste A, Cerrada E, Rodriguez Yoldi MJ. 2017. Colorectal carcinoma: a general overview and future perspectives in colorectal cancer. *Int. J. Mol. Sci.* 18(1):197
60. McGovern DP, Kugathasan S, Cho JH. 2015. Genetics of inflammatory bowel diseases. *Gastroenterology* 149:1163–76.e2

61. Meng M, Zhong K, Jiang T, Liu Z, Kwan HY, Su T. 2021. The current understanding on the impact of KRAS on colorectal cancer. *Biomed. Pharmacother.* 140:111717
62. Mihindukulasuriya KA, Mars RAT, Johnson AJ, Ward T, Priya S, et al. 2021. Multi-omics analyses show disease, diet, and transcriptome interactions with the virome. *Gastroenterology* 161:1194–207.e8
63. Minot S, Sinha R, Chen J, Li H, Keilbaugh SA, et al. 2011. The human gut virome: inter-individual variation and dynamic response to diet. *Genome Res.* 21:1616–25
64. Mirsepasi-Lauridsen HC, Vallance BA, Krogfelt KA, Petersen AM. 2019. *Escherichia coli* pathobionts associated with inflammatory bowel disease. 32(2):e00060-18
65. Monaghan TM, Sloan TJ, Stockdale SR, Blanchard AM, Emes RD, et al. 2020. Metagenomics reveals impact of geography and acute diarrheal disease on the Central Indian human gut microbiome. *Gut Microbes* 12:1752605
66. Nakatsu G, Zhou H, Wu WKK, Wong SH, Coker OO, et al. 2018. Alterations in enteric virome are associated with colorectal cancer and survival outcomes. *Gastroenterology* 155:529–41.e5
67. Nayfach S, Páez-Espino D, Call L, Low SJ, Sberro H, et al. 2021. Metagenomic compendium of 189,680 DNA viruses from the human gut microbiome. *Nat. Microbiol.* 6:960–70
68. Neri U, Wolf YI, Roux S, Camargo AP, Lee B, et al. 2022. Expansion of the global RNA virome reveals diverse clades of bacteriophages. *Cell* 185(21):4023–37.e18
69. Nishijima S, Nagata N, Kiguchi Y, Kojima Y, Miyoshi-Akiyama T, et al. 2022. Extensive gut virome variation and its associations with host and environmental factors in a population-level cohort. *Nat. Commun.* 13:5252
70. Norman JM, Handley SA, Baldridge MT, Droit L, Liu CY, et al. 2015. Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell* 160:447–60
71. Odamaki T, Kato K, Sugahara H, Hashikura N, Takahashi S, et al. 2016. Age-related changes in gut microbiota composition from newborn to centenarian: a cross-sectional study. *BMC Microbiol.* 16:90
72. Oh JH, Alexander LM, Pan M, Schueler KL, Keller MP, et al. 2019. Dietary fructose and microbiota-derived short-chain fatty acids promote bacteriophage production in the gut symbiont *Lactobacillus reuteri*. *Cell Host Microbe* 25:273–84.e6
73. Paez-Espino D, Zhou J, Roux S, Nayfach S, Pavlopoulos GA, et al. 2019. Diversity, evolution, and classification of virophages uncovered through global metagenomics. *Microbiome* 7:157
74. Pérez-Brocal V, García-López R, Vázquez-Castellanos JF, Nos P, Beltrán B, et al. 2013. Study of the viral and microbial communities associated with Crohn's disease: a metagenomic approach. *Clin. Transl. Gastroenterol.* 4:e36
75. Roux S, Adriaenssens EM, Dutilh BE, Koonin EV, Kropinski AM, et al. 2019. Minimum Information about an Uncultivated Virus Genome (MIUViG). *Nat. Biotechnol.* 37:29–37
76. Roux S, Hallam SJ, Woyke T, Sullivan MB. 2015. Viral dark matter and virus-host interactions resolved from publicly available microbial genomes. *eLife* 4:e08490
77. Schulfer A, Santiago-Rodriguez TM, Ly M, Borin JM, Chopyk J, et al. 2020. Fecal viral community responses to high-fat diet in mice. *mSphere* 5(1):e00833-19
78. Secor PR, Sweere JM, Michaels LA, Malkovskiy AV, Lazzareschi D, et al. 2015. Filamentous bacteriophage promote biofilm assembly and function. *Cell Host Microbe* 18:549–59
79. Sender R, Fuchs S, Milo R. 2016. Revised estimates for the number of human and bacteria cells in the body. *PLOS Biol.* 14:e1002533
80. Shen S, Huo D, Ma C, Jiang S, Zhang J. 2021. Expanding the colorectal cancer biomarkers based on the human gut phageome. *Microbiol. Spectr.* 9:e0009021
81. Shkoporov AN, Clooney AG, Sutton TDS, Ryan FJ, Daly KM, et al. 2019. The human gut virome is highly diverse, stable, and individual specific. *Cell Host Microbe* 26:527–41.e5
82. Shkoporov AN, Hill C. 2019. Bacteriophages of the human gut: the "known unknown" of the microbiome. *Cell Host Microbe* 25:195–209
83. Shkoporov AN, Khokhlova EV, Stephens N, Hueston C, Seymour S, et al. 2021. Long-term persistence of crAss-like phage crAss001 is associated with phase variation in *Bacteroides intestinalis*. *BMC Biol.* 19:163
84. Shkoporov AN, Ryan FJ, Draper LA, Forde A, Stockdale SR, et al. 2018. Reproducible protocols for metagenomic analysis of human faecal phageomes. *Microbiome* 6:68

85. Shkoporov AN, Turkington CJ, Hill C. 2022. Mutualistic interplay between bacteriophages and bacteria in the human gut. *Nat. Rev. Microbiol.* 20(12):737–49
86. Sinha A, Li Y, Mirzaei MK, Shamash M, Samadfam R, et al. 2022. Transplantation of bacteriophages from ulcerative colitis patients shifts the gut bacteriome and exacerbates the severity of DSS colitis. *Microbiome* 10:105
87. Siringan P, Connerton PL, Cummings NJ, Connerton IF. 2014. Alternative bacteriophage life cycles: the carrier state of *Campylobacter jejuni*. *Open. Biol.* 4:130200
88. Smith L, Goldobina E, Govi B, Shkoporov AN. 2023. Bacteriophages of the order *Crassvirales*: What do we currently know about this keystone component of the human gut virome? *Biomolecules* 13(4):584
89. Sutton TDS, Clooney AG, Hill C. 2020. Giant oversights in the human gut virome. *Gut* 69(7):1357–58
90. Sutton TDS, Clooney AG, Ryan FJ, Ross RP, Hill C. 2019. Choice of assembly software has a critical impact on virome characterisation. *Microbiome* 7:12
91. Sutton TDS, Hill C. 2019. Gut bacteriophage: current understanding and challenges. *Front. Endocrinol.* 10:784
92. Taboada B, Morán P, Serrano-Vázquez A, Iša P, Rojas-Velázquez L, et al. 2021. The gut virome of healthy children during the first year of life is diverse and dynamic. *PLOS ONE* 16:e0240958
93. Tisza MJ, Buck CB. 2021. A catalog of tens of thousands of viruses from human metagenomes reveals hidden associations with chronic diseases. *PNAS* 118(23):e2023202118
94. Tomofuji Y, Kishikawa T, Maeda Y, Ogawa K, Nii T, et al. 2022. Whole gut virome analysis of 476 Japanese revealed a link between phage and autoimmune disease. *Ann. Rheum. Dis.* 81:278–88
95. Tomova A, Bukovsky I, Rembert E, Yonas W, Alwarith J, et al. 2019. The effects of vegetarian and vegan diets on gut microbiota. *Front. Nutr.* 6:47
96. Turner D, Kropinski AM, Adriaenssens EM. 2021. A roadmap for genome-based phage taxonomy. *Viruses* 13(3):506
97. Turner D, Shkoporov AN, Lood C, Millard AD, Dutilh BE, et al. 2023. Abolishment of morphology-based taxa and change to binomial species names: 2022 taxonomy update of the ICTV bacterial viruses subcommittee. *Arch. Virol.* 168:74
98. Van Espen L, Bak EG, Beller L, Close L, Deboutte W, et al. 2021. A previously undescribed highly prevalent phage identified in a Danish enteric virome catalog. *mSystems* 6:e0038221
99. Wagner J, Maksimovic J, Farries G, Sim WH, Bishop RF, et al. 2013. Bacteriophages in gut samples from pediatric Crohn's disease patients: metagenomic analysis using 454 pyrosequencing. *Inflamm. Bowel Dis.* 19:1598–608
100. Yang K, Niu J, Zuo T, Sun Y, Xu Z, et al. 2021. Alterations in the gut virome in obesity and type 2 diabetes mellitus. *Gastroenterology* 161:1257–69.e13
101. Yutin N, Benler S, Shmakov SA, Wolf YI, Tolstoy I, et al. 2021. Analysis of metagenome-assembled viral genomes from the human gut reveals diverse putative CrAss-like phages with unique genomic features. *Nat. Commun.* 12:1044
102. Yutin N, Makarova KS, Gussow AB, Krupovic M, Segall A, et al. 2018. Discovery of an expansive bacteriophage family that includes the most abundant viruses from the human gut. *Nat. Microbiol.* 3:38–46
103. Zhao M, Feng R, Ben-Horin S, Zhuang X, Tian Z, et al. 2022. Systematic review with meta-analysis: environmental and dietary differences of inflammatory bowel disease in Eastern and Western populations. *Aliment. Pharmacol. Ther.* 55:266–76
104. Zuo T, Liu Q, Zhang F, Yeoh YK, Wan Y, et al. 2021. Temporal landscape of human gut RNA and DNA virome in SARS-CoV-2 infection and severity. *Microbiome* 9:91
105. Zuo T, Lu XJ, Zhang Y, Cheung CP, Lam S, et al. 2019. Gut mucosal virome alterations in ulcerative colitis. *Gut* 68:1169–79
106. Zuo T, Sun Y, Wan Y, Yeoh YK, Zhang F, et al. 2020. Human-gut-DNA virome variations across geography, ethnicity, and urbanization. *Cell Host Microbe* 28:741–51.e4
107. Zuo W, Michail S, Sun F. 2022. Metagenomic analyses of multiple gut datasets revealed the association of phage signatures in colorectal cancer. *Front. Cell Infect. Microbiol.* 12:918010