The human gut virome: a multifaceted majority

Lesley A. Ogilvie[1,2]* and Brian V. Jones[1,3]

[1] School of Pharmacy and Biomolecular Sciences, University of Brighton, Brighton, UK, [2] Alacris Theranostics GmbH, Berlin, Germany, [3] Queen Victoria Hospital NHS Foundation Trust, East Grinstead, UK

Here, we outline our current understanding of the human gut virome, in particular the phage component of this ecosystem, highlighting progress, and challenges in viral discovery in this arena. We reveal how developments in high-throughput sequencing technologies and associated data analysis methodologies are helping to illuminate this abundant 'biological dark matter.' Current evidence suggests that the human gut virome is a highly individual but temporally stable collective, dominated by phages exhibiting a temperate lifestyle. This viral community also appears to encode a surprisingly rich functional repertoire that confers a range of attributes to their bacterial hosts, ranging from bacterial virulence and pathogenesis to maintaining host–microbiome stability and community resilience. Despite the significant advances in our understanding of the gut virome in recent years, it is clear that we remain in a period of discovery and revelation, as new methods and technologies begin to provide deeper understanding of the inherent ecological characteristics of this viral ecosystem. As our understanding increases, the nature of the multi-partite interactions occurring between host and microbiome will become clearer, helping us to more rationally define the concepts and principles that will underpin approaches to using human gut virome components for medical or biotechnological applications.

Keywords: bacteriophage, human gut microbiome, metagenomics, virus-like particles, phage-encoded functions, dysbiosis

Introduction

In recent years it has become apparent that the microbes resident in and on us play a significant role in our health and well-being. These microbial residents have been shown to provide a range of essential functions, from helping us to digest food to conditioning our immune system, and providing protection from invading pathogens (Rakoff-Nahoum et al., 2004; Bäckhed et al., 2005; Ley et al., 2006; O'Hara and Shanahan, 2006; Zoetendal et al., 2006; Sokol et al., 2008). The advent of high-throughput sequencing technologies is enabling further insights into this complex human–microbe relationship, revealing significant associations between microbial ecosystem shifts and disease (termed dysbiosis), and bringing to light a diverse and abundant retinue of viruses intimately associated with the human gut microbiome. This human gut virome may be defined as the total population of viruses [or virus like particles (VLPs)] associated with the underlying gut microbial community. In keeping with the dominance of bacteria in the gut microbiome, the gut virome appears to be predominated by prokaryotic viruses (bacteriophage, or phage; Breitbart et al., 2003; Reyes et al., 2010; Minot et al., 2011). These encompass both DNA and RNA viruses that infect bacteria, but the latter are rare members of the phage family overall, and studies

to date suggest these are accordingly a very minor component of the gut virome (Breitbart et al., 2003; Reyes et al., 2010; Minot et al., 2011).

Since the independent discovery of phages by Twort (1915) and d'Herelle (1917), the potential therapeutic and biotechnological application of these bacterial viruses has been a source of intense research interest. Phages are the most abundant infectious agents on the planet ($>10^{31}$, extrapolated from Whitman et al., 1998), outnumbering their bacterial hosts, on average by 10:1 in most habitats studied (Wommack et al., 1999). Through their ability to lyse and kill host bacteria, phages play a crucial role in modulating bacterial community structure and function (Gorski et al., 2003; Brüssow et al., 2004; Suttle, 2005a,b, 2007; Rohwer and Thurber, 2009). This is perhaps most profoundly exemplified by the impact of phages on global biogeochemical cycles, and potentially the climate, exerted through effects on marine microbial ecology (Fuhrman, 1999; Breitbart, 2012; Brüssow, 2013).

Phages may also influence community function through the facilitation of gene transfer between strains and species (transduction), or themselves encode accessory functions of benefit to host species. Key functional attributes conferred from phages to their bacterial hosts include toxin synthesis, production of virulence factors, as well as genes that may provide metabolic flexibility (Fuhrman, 1999; Wommack and Colwell, 2000; Brüssow et al., 2004; Suttle, 2007). However, phage-host relationships are not static but dynamic co-evolutionary interactions, which play an important role in evolution of their bacterial hosts (Paterson et al., 2010). This repertoire of attributes and influences on bacterial hosts place phages as a potent force driving ecological functioning and evolutionary change in the prokaryotic world (Koskella and Brockhurst, 2014).

Phage (and virome) research is currently entering a renaissance, driven by an era of technological advancement that is seeing a shift from reductionist thinking. This advancement has made feasible investigation of complex microbial communities (such as the human gut microbiome) tractable on a systems level; a paradigm now increasingly extended to the viral components associated with such ecosystems (Reyes et al., 2012). Developments in high-throughput sequencing technologies and associated bioinformatics analyses are starting to provide crucial insights into the unique ecological characteristics of the human gut virome, shedding light on the structure and functioning of this viral collective. Recent work has revealed a surprisingly rich functional repertoire is encoded by the gut virome, with phages conferring a range of beneficial traits to their bacterial hosts that help maintain community stability, and afford resilience to invasion or disruption (Breitbart et al., 2003, 2008; Zhang et al., 2006; Reyes et al., 2010; Colomer-Lluch et al., 2011; Kim et al., 2011; Minot et al., 2011, 2012a,b; Modi et al., 2013; Ogilvie et al., 2013). Given this renewed interest and realization that human gut phages may play a significant role in shaping the development and functional outputs of host microbiomes, their potential for application in novel diagnostic, therapeutic, and biotechnological applications is now of considerable interest.

Here, we review recent insights into the human gut virome, focussing on the dominant phage component associated with this ecosystem, and examine currently available approaches to access, dissect, and elucidate the role of this viral collective. In doing so, we highlight the challenges to viral discovery and future routes of investigation.

Life within the Human Gut

One of the most densely populated areas of the human body is the gastrointestinal tract (GIT), which provides a heterogeneous and expansive surface area (>200 m^2) for microbial life (Hooper and Macpherson, 2010). The human gut is estimated to contain between 30 and 400 trillion micro-organisms (Suau et al., 1999), drawn from varied and abundant bacterial hosts that can support a correspondingly rich and varied phage population. Numbers of bacteria vary across the length of the GIT, with an enrichment of different bacterial types at different sites (Ahmed et al., 2007; Zhang et al., 2014). Corresponding studies of the variation in associated phage/virome populations have yet to be conducted in the human gut, although spatial analyses of the swine gut have revealed a differential distribution with increased abundance of phage encoded genes in the ileum, coinciding with reduced bacterial numbers (Looft et al., 2014). These observations have been linked to gut physiology. Due to the cycle of feast and famine likely occurring in the ileum (as food is ingested and processed), and the role of phages in nutrient release through lysis of host bacteria, there could be fitness costs associated with phage resistance in ileal bacteria (Looft et al., 2014). The physical and physiochemical heterogeneity is further enhanced by a thick mucosal layer that separates the intestinal epithelium from the lumen. This highly dynamic mucosal layer is in constant flux, acting both as a protective barrier to invading pathogens and residence for commensal bacteria and their attendant viral collective (Barr et al., 2013a,b). The physical conditions likely created by this mucous layer will undoubtedly influence the nature of the interactions which occur between the human host, bacteria, and phage. Indeed, enrichment of phage numbers within mucosal environments as well as elevated ratios of phages to bacteria have been reported (Barr et al., 2013a), as well as the potential role of these phage and the mucosal environment in the establishment of a non-host derived phage based innate immunity (Barr et al., 2013a).

Human Gut Virome Characteristics

In many environments phages have been found to outnumber their bacterial hosts by an order of magnitude (Brüssow and Hendrix, 2002; Suttle, 2005a). In the human gut, however, phages are believed to exist at levels comparable to their bacterial hosts. Based on microscopy counts, bacteria appear typically at $\sim 10^9$/g feces and VLPs $\sim 10^8$–10^9/g (Kim et al., 2011), with most VLPs originating from phages (Reyes et al., 2010; Minot et al., 2011). VLP-derived estimates of phage diversity, in terms of phylotypes, also reflect that of bacterial hosts, with estimated ratios of 1:1 (phage, bacteria; Reyes et al., 2010; Minot et al., 2011). Similar to their bacterial hosts, phages have been found

to accumulate on gut mucosal surfaces and within mucous, but here are found at much higher levels with phage, bacteria ratios of approximately 20:1 within the mouse intestinal mucosa (Barr et al., 2013a).

In contrast to non-host associated ecosystems such as aquatic environments, predator-prey interactions or 'kill-the-winner' phage-host dynamics (Brüssow and Hendrix, 2002; Thingstad et al., 2008), are seemingly lacking in the human gut microbiome (Reyes et al., 2010). Instead, most of the dominant virotypes detected in the gut ecosystem show evidence for a temperate lifestyle (in which phages integrate into host chromosomes or exist as quiescent episomal elements at the expense of lytic replication), as indicated by the frequency of integrase genes and other genetic features (Breitbart et al., 2003; Reyes et al., 2010; Minot et al., 2011). The genetic symbiosis established by temperate phage, persisting as prophage within host cells, is important for genetic exchange between bacterial hosts, alteration of host phenotypes *via* lysogenic conversion (Brüssow et al., 2004; Reyes et al., 2010), which in turn impacts on bacterial host fitness as well as human gut microbial dynamics (Thingstad et al., 2008; Duerkop et al., 2012). In this respect, the carriage of phages as dormant prophage may still have profound implications for the competitive abilities of bacterial hosts, potentially ensuring dominance of their intestinal niche in the presence of closely related competing strains (Thingstad et al., 2008; Duerkop et al., 2012).

These observations do not exclude the possibility that other phage life cycles and strategies exist. Intriguing observations within the oral virome (Abeles et al., 2014) have suggested that phage could also be lysing hosts at a constitutive low level, with longitudinal analyses of the oral viromes of eight human subjects consistently identifying specific phages at all time points over a 60-days period and in all subjects sampled (the specific phage detected varying between individuals); while observations of phages accumulating within mucosal environments (Barr et al., 2013a) open up the possibility of a phage survival strategy, with the mucosa facilitating phage persistence until a suitable host is encountered (Edlund et al., 2015). Whether the same holds true for the human gut is a question still to be answered.

Although human gut phages are often found to have high specificity to host bacterial species and even strains *in vitro* (e.g., Weinbauer, 2004; Ogilvie et al., 2012b), this phenomena may only be an artifact of the laboratory setting (Chibani-Chennoufi et al., 2004). Observations of phage host range expansion following antibiotic exposure within the mouse gut (Modi et al., 2013) highlight the adaptive capacity and potential genetic plasticity of phages (Chibani-Chennoufi et al., 2004), but nevertheless *in vivo* confirmation of phage host range and specificity within the human gut is severely lacking.

The human gut virome is also seemingly characterized by a high degree of inter-personal variation (much higher than the associated gut microbiome) and remarkable temporal stability in structure (Reyes et al., 2010; Minot et al., 2011, 2013). Reyes et al. (2010) report retention of >95% of viral genotypes within one individual, with minimal changes in relative abundances, over a 1 year period. Similarly, Minot et al. (2011) report persistence of 80% of virotypes within an individual over a 2.5 years period.

The observed high level of inter-individual variation has been linked to variation in the bacterial component of the human gut microbiome carried by individuals, which has been found to vary at the level of species and strain (Huttenhower et al., 2012), persistence of a small portion of the global virome within each individual, and also the ability of viral populations (in particular lytic phage) to undergo rapid evolution to form new "species" of phages (Minot et al., 2013). Greater similarity between viromes is found, however, when individuals eat the same diet (Minot et al., 2011), likely reflecting dietary induced convergence also found in their bacterial hosts (Wu et al., 2011). Moreover, recent studies, mainly based on bioinformatics-based interrogation of whole community metagenomes, are demonstrating that there may be a higher level of virotype conservation between individuals than previously thought (Stern et al., 2012; Ogilvie et al., 2013; Dutilh et al., 2014).

Many of these assumptions have, however, been made based on the analysis of a relatively limited number of individuals. In comparison, many of the studies investigating the diversity of the human gut bacterial component of the gut ecosystem, i.e., based on the analysis of 16S rRNA sequences, involve study numbers of generally an order of magnitude or more. In the future, a more detailed interrogation of the human gut virome over a wider range of temporal and spatial scales and comprising larger cohorts is required to fully delineate the characteristics of this viral ecosystem.

Accessing and Analyzing the Human Gut Virome – Challenges and Solutions

Our current state of knowledge of the human gut virome has been informed by a combination of traditional culture-based techniques, high-resolution microscopy, and metagenomic analyses. Taking an analogous trajectory to that of characterizing the bacterial component of human gut microbiome, approaches to sampling the associated complex viral communities populating this ecosystem have moved from reliance on more traditional culture and microscopy-based methodologies, toward metagenomic analysis of nucleic acids by next generation sequencing technologies. **Table 1** provides an overview of current approaches commonly used for accessing and understanding the human gut virome.

Culture-based techniques involving the isolation of phages from the environment using a specific bacterial host have informed our basic understanding of phages for decades and will continue to be an essential tool for human gut viral discovery. Indeed, our present knowledge of phage-host interactions, functional repertoire, and phage ecology is largely underpinned by numerous detailed *in vitro* analyses based on conventional plaque assays used to isolate phages and associated virology. Such techniques are inherently limited by the inability to culture >99% of bacterial species found in many environments (Amann et al., 1995; Rappé and Giovannoni, 2003), however, ongoing efforts are substantially increasing representation of culturable gut inhabitants (Rajilić-Stojanović and de Vos, 2014). There is also uncertainty regarding the degree to which observations

TABLE 1 | Benefits and limitations of commonly used methods for accessing the human gut virome.

Method	Description	Benefits	Limitations
Culture-based			
– Plaque-based assays	Isolation of phages from a biological sample using a cultured bacterial host	• Essential tool that has underpinned our understanding of phage-host interactions • Bacterial host is known • Host–phage dynamics can be investigated in real time • Inexpensive • Accessible for most laboratories • Amenable to testing hypotheses	• Culture dependent • Limited range of bacterial hosts, due to inability to culture >99% of bacteria • Recapitulation of *in vivo* observations *in vitro* is questionable
Microscopy			
– Transmission electron microscopy (TEM)	Visualization of viruses following negative staining	• Culture independent • Provides *in situ* high resolution structural visualization (TEM) of viruses • Provides information of morphological diversity of viral/phage types • Useful as supplement to other approaches	• No functional information • Little or no community level insight • No host range information unless used to view purified phages propagated on a culturable bacterial host • Requires high virus like particle/phage titres for visualization • Phages with unusual capsid structures, e.g., Inoviridiae, may be difficult to identify
– Fluorescence	Visualization of viruses following staining by DNA fluorochromes	• Culture independent • Quantitative • Rapid • Cost effective • Broad overview of viral numbers	• As for TEM, but additionally, information on viral structure and morphology is not provided, with information largely restricted to viral numbers
PCR-based			
– PCR/quantitative PCR	Amplification of DNA segments using primers targeting a specific genomic section	• Culture independent • Quantitative	• Lack of conserved genes within all phages akin to the 16S rRNA gene in bacteria • View of viral/phage abundance and diversity limited to available sequence data from which primers are designed
Metagenomics			
– Large/small insert metagenomic libraries	Isolation of nucleic acids from virus like particles followed by direct cloning in surrogate host species	• Culture independent • Can provide access to accessory genes and potential assessment of function/impact on host bacteria • Can provide community-level insights	• Potential bias in inventory gained as focuses on free viral particles (i.e., actively replicating phage) • Laborious and expensive, with insights limited by scale of libraries that can be generated and screened • Biases due to difficulties in cloning phage DNA and maintaining in surrogate host may be apparent • Use of single surrogate host may limit heterologous gene expression and therefore utility in functional evaluation of phage genes • Loss of host range information
– Virus-like particle libraries	Isolation of virus like particles from biological samples using centrifugation and/or filtration methods, extraction, and amplification of encapsulated DNA followed by high-throughput sequencing	• Culture independent • In-depth inventory of composition and functional repertoire of viruses within biological samples • Assessment of viral diversity and lifestyles within the human gut • Community-level assessment	• Potential bias in inventory gained as focuses on free viral particles (i.e., actively replicating phage) • Viral DNA recovered can be very low therefore amplification techniques used that may bias community snapshot gained • No host information • Large fraction of recovered sequences are 'unknown' • Potential contamination of VLP libraries with bacterial DNA • Relatively expensive for routine/repeated analyses • Difficult to test hypotheses generated • Loss of host range information
– Whole community shotgun sequencing	Extraction of total DNA from a biological sample, followed by library preparation and high-throughput sequencing	• Viral DNA is significant fraction of extracted DNA • Full microbial community snapshot gained • Community level inventory • Structural and functional information provided	• Difficult to distinguish between bacterial and viral sequences • Identification of host species remains challenging • Potential bias in inventory gained as focuses on 'quiescent' fraction of virome • Relatively expensive for routine/repeated analyses • Difficult to test hypotheses generated

generated under highly controlled, standardized, and essentially artificial laboratory conditions reflect the situation *in vivo*. This is particularly true for understanding the functioning of complex communities such as the gut microbiome where, although there are some good systems available that replicate some of the features of this system, it is still difficult to develop realistic and representative *in vitro* models.

High resolution microscopy analysis techniques such as transmission electron microscopy (TEM) have also proved essential for the study of phage structure and identification of phage types with distinct morphologies. Prior to electron microscopy the structure and basic nature of viruses in general was a complete unknown, with their existence known only through observations that for some diseases, cell free tissue extracts remained capable of causing the original disease. This provided the origin for the term virus, which was derived from the latin word describing a poisonous or venomous substance. Electron microscopy has also provided early insight into the composition of the human gut viral community based on phage structure and morphology, which continue to be important parameters in viral classification systems. Such studies revealed tailed phages (Order *Caudovirales*) to be prevalent (Flewett et al., 1974; Letarov and Kulikov, 2009); observations corroborated and expanded by more recent metagenomic-based catalogs, showing a dominance of tailed, doubled stranded DNA viruses of the order *Caudovirales* (*Siphoviridae, Myoviridae, Podoviridae*) alongside the tail-less single stranded DNA viruses (*Microviridae*; Reyes et al., 2010; Kim et al., 2011).

Recent rapid progress in sequencing technologies and associated bioinformatics methodologies has enabled a more in-depth view of the structure and functioning of these viral communities. Development of methods to isolate VLPs from fecal material using a combination of filtration and density gradient ultra-centrifugation techniques, followed by extraction and amplification of viral nucleic acids and subsequent sequencing (Thurber et al., 2009; Reyes et al., 2012), have provided unprecedented glimpses into the viral fraction of the human gut ecosystem, providing the first in-depth inventory of the composition and functional repertoire of this collective (Breitbart et al., 2003; Reyes et al., 2010; Kim et al., 2011; Minot et al., 2011).
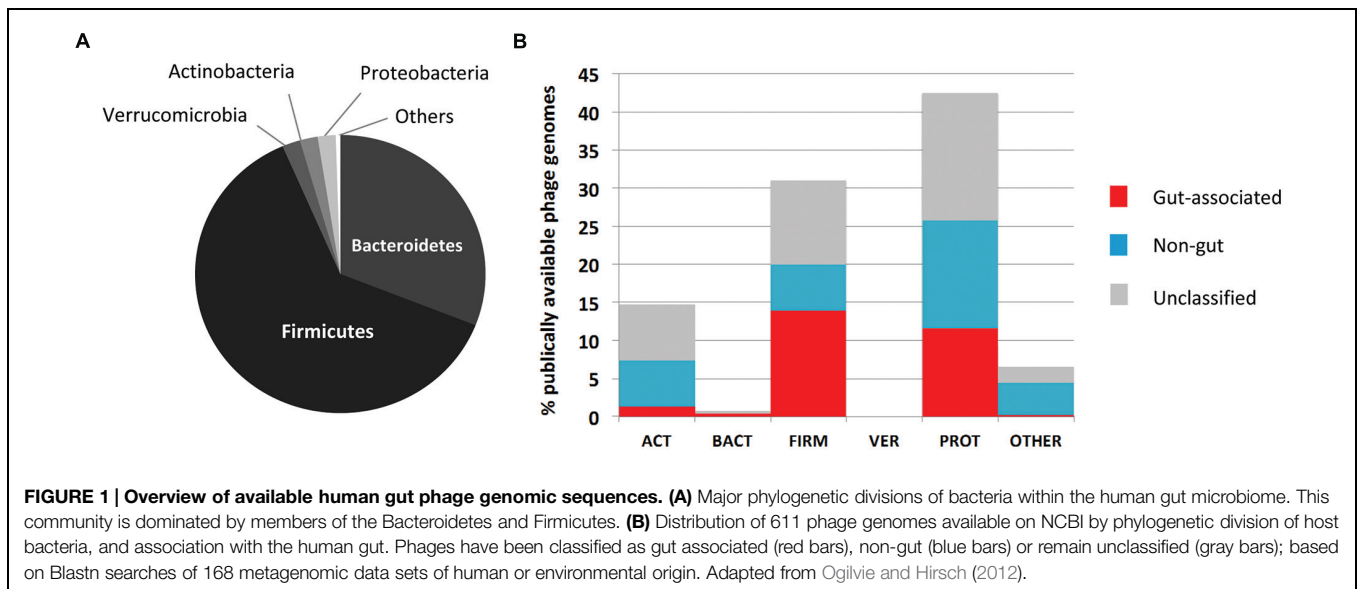
In keeping with studies of the bacterial fraction, such metagenomic approaches have now become the mainstay of viral ecology and are set to usher in a new era of viral ecogenomics akin the renaissance recently enjoyed by microbial ecology. There remain, however, a number of challenges associated with the use and interpretation of results generated using metagenomic techniques. Inherent in the extraction of VLPs is a focus on the analysis of free and therefore actively replicating phage particles present at the time of sampling, restricting access to the quiescent fraction. In addition, amplification procedures implemented due to low levels of extracted DNA, potentially exclude the analysis of certain phage types (Angly et al., 2006), biasing the snapshot of the virome gained from these methods. However, perhaps most important for community level ecogenomic investigations is the lack of host range information afforded by metagenomic approaches. Although the culture-independent nature of these

methods is a key advantage, this effectively divorces the resulting sequence data from bacterial host-range information, necessitating the use of additional bioinformatic approaches to infer this indirectly.

Our understanding of the functional and phylogenetic make-up of the human gut phage/viral gene space has largely been based on alignment-based techniques such as the Blast suite of algorithms (Altschul et al., 1990), exploiting homologies between sequences at the nucleotide and amino acid levels. These methods have also highlighted the vast level of novel gene content encoded by phage genomes, evident in the lack of homology of many phage genes to existing database entries (Breitbart et al., 2003; Reyes et al., 2010; Minot et al., 2011). Such novelty poses a significant challenge to efforts seeking to understand the basic structure and function of viral communities, and is compounded further by the dearth of well-characterized phage reference genome sequences derived from those viruses established to infect key members of the gut microbiome (**Figure 1**). This is exemplified by the lack of genomic data on phages infecting *Bacteroides* spp., prominent human gut members and implicated in both the onset of and protection against inflammatory-mediated disease (Wexler, 2007; Mazmanian et al., 2008), with only two representative genomes (e.g., Hawkins et al., 2008; Ogilvie et al., 2012b) currently (March 2015) deposited in global sequence databanks. On a more general level, estimates suggest less than 0.001% of the predicted phage diversity is represented in global sequence databanks (Bibby, 2014). In contrast, the interpretation of metagenomic data derived from the bacterial components of the gut ecosystem has benefitted greatly from access to an ever expanding catalog of complete genome sequence data from diverse organisms and habitats.

Evolutionary and ecological insight into the gut virome has also been hampered by the lack of a common phylogenetic anchor in phages (akin to the 16S rRNA gene in bacteria), enabling detailed studies of the phylogenetic relationships existing between ecosystem members. Although conserved phage genes do exist, such as the terminase gene (Casjens, 2003), the significant mosaicism exhibited by phage genomes obscure any robust phylogenetic signal.

Whole-community metagenomes, generated from DNA directly extracted from fecal material and containing significant fractions of viral DNA (4–17%, Qin et al., 2010; Minot et al., 2011), are now being considered as a valuable resource for the analysis of phage communities that when placed in tandem with VLP-derived datasets have the potential to provide a more cohesive overview of the phage component of ecosystems, including the human gut. Despite these potential benefits, both VLP-derived and whole community metagenomic datasets present a number of challenges for analysis of the data from a viral perspective. In particular, distinguishing between viral and bacterial sequence reads, allocating reads to distinct genomes, and assigning host species to identified phage. As noted above, these challenges are further compounded not just by the diversity of the virosphere, but also the lack of complete and well-characterized phage reference genomes with established host range and functionality.

FIGURE 1 | Overview of available human gut phage genomic sequences. (A) Major phylogenetic divisions of bacteria within the human gut microbiome. This community is dominated by members of the Bacteroidetes and Firmicutes. (B) Distribution of 611 phage genomes available on NCBI by phylogenetic division of host bacteria, and association with the human gut. Phages have been classified as gut associated (red bars), non-gut (blue bars) or remain unclassified (gray bars); based on Blastn searches of 168 metagenomic data sets of human or environmental origin. Adapted from Ogilvie and Hirsch (2012).

To overcome some of these hurdles a range of inventive and insightful viral-orientated bioinformatics tools, pipelines, and approaches have emerged in recent years, which are aimed at standardizing and improvement of assembly, annotation, and comparative analysis of viral metagenomes. These include VIROME (Viral Informatics Resource for Metagenome Exploration; Wommack et al., 2012), MetaVir (Roux et al., 2014), iVireons (Seguritan et al., 2012), and PHACCS (PHAge Communities from Contig Spectrum; Angly et al., 2005). Application of these viral/phage-orientated analysis methodologies are starting to heighten our understanding of this novel gene space.

Stern et al. (2012) exploited the anti-phage immune system of bacteria, the CRISPR (clustered regularly interspaced short palindromic repeats)/Cas (CRISPR-associated) system, to identify phage sequences within human gut microbial metagenomic datasets. The CRISPR/Cas system represents a mechanism of bacterial acquired adaptive immunity against phage, based on a genome-based catalog of previous infections. The CRISPR system comprises short repeated sequences that are separated by hypervariable CRISPR spacers of between 24 and 50 bp. Bacteria incorporate DNA of the infecting phage into these spacers, effectively establishing a genome-based phage resistance catalog, which when transcribed into small RNA and in tandem with formation of a Cas protein complex, block the replication of similar phages (Sorek et al., 2008). The CRISPR spacers identified by Stern et al. (2012) within raw sequencing reads of 124 whole community metagenomes (Qin et al., 2010), were used to probe assembled sequence contigs; this resulted in detection of approximately 1000 partial and complete phage genomes, which appeared in multiple individuals of diverse geographic origin (Stern et al., 2012). Bacterial hosts for some of these phage were identified and insight into patterns of phage-bacteria co-existence was gained. Results which pointed to the existence of a reservoir of phage frequently associated with the gut microbiome, but no obvious evidence for the existence of similarities in gut

viromes, or enterotypes, as posited for the microbial component (Arumugam et al., 2011).

Taking advantage of similarities in global nucleotide usage patterns, or the genome signature, arising between phages infecting the same or related host bacterial species (Pride et al., 2003, 2012), a genome-signature based approach has also been developed for the phage-orientated dissection of conventional whole-community metagenomes (Ogilvie et al., 2013). Comparison of tetranucleotide usage profiles (patterns of four nucleotides) exhibited by a well-characterized set of phage genomes with known host-range, were compared to 139 human gut microbial metagenomic datasets using TETRA (Teeling et al., 2004). Sequences exhibiting similar tetranucleotide frequency patterns were identified, retrieved and functionally profiled to identify sequences of phage origin, infer host range and provide insights into phage lifestyles within the human gut. Using this approach a portion of the human gut virome poorly represented in VLP-derived metagenomes was revealed, illuminating phylogenetically distinct, predominantly temperate, phage genomes encoding a range of functions relevant to human health (Ogilvie et al., 2013). Furthermore, shared patterns of distribution in multiple individuals were identified indicating the existence of putative viral enterotypes (Ogilvie et al., 2013), and providing continuity with the ever expanding knowledge base on bacterial constituents of the human gut microbiome (Arumugam et al., 2011).

Most recently, application of a *cross-assembly* method (Dutilh et al., 2012, 2014) – a bioinformatics approach that assumes sequences repeatedly found in different metagenomes are most likely to be part of the same genome – for the reanalysis of 12 existing human gut virome datasets (Reyes et al., 2010), revealed the existence of a highly abundant 92 kb phage in multiple individuals (Dutilh et al., 2014). Interestingly, these most recent publications demonstrating alternative approaches to the re-analysis of human gut microbiomes and viromes (Stern et al., 2012; Ogilvie et al., 2013; Dutilh et al., 2014), challenge the

notion that our gut viromes are highly individualized due to the perceived fast-paced evolutionary race with their bacterial hosts, and are perhaps surprising in light of the vast genetic diversity reported in many ecosystems.

While these collective observations are seemingly contradictory to studies focusing on the generation and analysis of VLP based metagenomes, it should be noted that they are based on using datasets generated using distinct approaches and therefore may afford distinct insights into the gut virome not possible with VLP based datasets. Unlike the bacteria-centric metagenomes used in the studies noted above (Ogilvie et al., 2012a; Stern et al., 2012; Dutilh et al., 2014), VLP-based libraries are focused instead on actively replicating "free" phage at the time of sampling. In contrast analysis of whole community metagenomes provide access to the quiescent fraction, i.e., phages that may be inactive or dormant due to lysogenisation and poorly represented or absent from VLP-derived datasets, as noted in genome signature based dissection of gut metagenomes (Ogilvie et al., 2013). Such phages may nevertheless make important functional contributions to the community in terms of providing bacterial hosts with accessory functions, and potentially also resistance to infection by related phage types. Such observations highlight the need for complementary approaches to accessing the human gut virome, if the entirety of this viral ecosystem is to be cataloged and understood.

The Human Gut Phage Functional Landscape

Within marine environments phage-encoded genes are known to make a significant impact on bacterial-mediated ecosystem functioning (Fuhrman, 1999; Suttle, 2005a,b), but the contribution of phages to the human gut microbial ecosystem is less well-established. The functional landscape of the human gut virome is, however, beginning to be mapped, revealing a surprisingly rich functional repertoire. Recent metagenomic surveys have revealed that gut associated phages encode genes that are generally beneficial for intestinal bacteria, including functions that help host bacteria adapt to their environment, relate to bacterial virulence, and functions that help maintain host microbiome stability and community resilience (Breitbart et al., 2003, 2008; Zhang et al., 2006; Reyes et al., 2010, 2012; Colomer-Lluch et al., 2011; Kim et al., 2011; Minot et al., 2011; Duerkop et al., 2012; Modi et al., 2013; Ogilvie et al., 2013).

Analysis of VLP libraries has revealed an abundance of genes involved in energy harvest, e.g., carbohydrate and amino acid metabolism (Reyes et al., 2010; Minot et al., 2011), including the BACON domain (Reyes et al., 2013; Dutilh et al., 2014), with potential roles in targeting glycoproteins and possibly host mucin (Mello et al., 2010; Reyes et al., 2013). These findings suggest that phages could be playing key roles in human host metabolism through indirect effects on their bacterial hosts. Whereas cryptic prophages (i.e., devoid of genes required for lytic development) have been shown to be essential to their bacterial hosts (*Escherichia coli*), enabling them to withstand adverse conditions, such as antibiotic, osmotic, oxidative and acid

stresses, increasing growth and influencing biofilm formation (Wang et al., 2010; Reyes et al., 2013).

Phages may also constitute a repository of 'beneficial' genes in the gut microbiome, safeguarding important activities and genetic information during adverse advents that lead to the disruption of the community, and its subsequent re-establishment. Global analysis of the gut virome in mice revealed that following exposure to antibiotics, gut phage populations were enriched for genes associated with antibiotic resistance, as well as those related to gut colonization, and growth and adaptation of their microbial hosts in this environment (Modi et al., 2013). In this scenario, the virome was theorized to provide a functional buffer of genes necessary for community recovery following exposure to agents such as antibiotics.

The phage-encoded antibiotic resistance gene pool within the human gut is significant and diverse (Reyes et al., 2010; Minot et al., 2011) and has been shown to be viable, mobile, and widely distributed within multiple individuals of diverse geographical origin (Modi et al., 2013; Ogilvie et al., 2013). These observations reinforce the notion that phages may play a more significant role in the emergence and spread of antibiotic resistance genes than previously thought (Ogilvie et al., 2013). Given that phages have been shown to move between biomes (Breitbart et al., 2004; Sano et al., 2004; Breitbart and Rohwer, 2005), and the recognition that phages are a reservoir of antibiotic resistance determinants, both within and out-with the human gut (Colomer-Lluch et al., 2011; Modi et al., 2013; Marti et al., 2014; Muniesa and Jofre, 2014), understanding of the factors that result in mobilization of phage-encoded functions may be crucial to develop strategies that limit the spread of, e.g., antibiotic resistance and the emergence of pathogenic bacterial strains.

Overlaying these more direct functional attributes or contributions of gut-associated phage, is their involvement in wider mechanisms of genetic transfer. An extensive network of gene exchange exists within the human microbiome as a whole; with the human gut estimated to exhibit one of the highest rates of transfer of all body sites analyzed (Smillie et al., 2011). The observed or estimated rate of gene exchange is undoubtedly influenced by the environmental and physiological parameters within the human gut milieu (Licht and Wilcks, 2005), but as part of the mobile metagenome (Jones and Marchesi, 2007; Jones, 2010; Ogilvie and Jones, 2013), phages likely play a key role in this network through the process of transduction (generalized or specialized); a process in which DNA is transferred from one bacterium to another. Given the abundance of phages in the gut microbiota, and that this community is essentially comprised of high numbers of closely related strains and species, it is conceivable such a mechanism could be active within the human gut, exemplified by the phage-mediated transduction of plasmids between *Lactococcus lactis* and *Streptococcus thermophilus* (Licht et al., 1999). Overall, the human gut phage functional repertoire, similar to other mobile genetic elements such as plasmids, aligns with challenges and activities important to life within the human gut (Jones and Marchesi, 2007; Jones, 2010; Ogilvie et al., 2012b; Ogilvie and Jones, 2013). Moreover, as the full extent of the role of phages as agents of horizontal transfer with the human gut is revealed, their impact on host functional outputs will

become clearer, with potential implications for human health and well-being.

An Emerging Role for Gut Phages in Health and Disease

In recent years there has been intense interest in characterizing the bacterial dysbiosis signature of disease, with alterations in the structure and functioning of the human gut microbiome being associated with many conditions from inflammatory bowel disease (IBD) and colorectal cancer, to diseases of the central nervous system (e.g., Sobhani et al., 2011; Clemente et al., 2012; Cryan and Dinan, 2012; Wang et al., 2012; Ohman et al., 2015). Nascent observations are now starting to link the associated retinue of phages with the pathogenesis of disorders associated with perturbation of the gut ecosystem (Furuse et al., 1983; Lepage et al., 2008; Pérez-Brocal et al., 2013; Wagner et al., 2013; Dinakaran et al., 2014; Norman et al., 2015).

Early culture-based studies revealed quantitative and qualitative differences in phages isolated from fecal samples derived from healthy individuals and those with leukemic and internal disorders, with generally higher levels of coliphages in patient samples as compared to healthy subjects (Furuse et al., 1983). Lepage et al. (2008) used epifluorescence microscopy to detail the abundance of VLPs within gut mucosal samples taken from healthy and diseased individuals, revealing an increased number of VLPs in individuals suffering from Crohn's disease (CD), observations supported more recently by other metagenomics based studies revealing an altered representation of phages in individuals with CD compared to healthy controls (Wagner et al., 2013) and increased enteric virome richness and decreased bacterial diversity in individuals suffering from CD and ulcerative colitis (Pérez-Brocal et al., 2013).

Deep sequencing of blood samples (plasma DNA) has also provided some evidence of altered phage taxonomic and abundance profiles associated with host bacterial species, within patients suffering cardiovascular disease (Dinakaran et al., 2014). In the latter study, altered phage composition is hypothesized to result from microbial translocation across the intestinal barrier, which becomes more permeable following cardiac surgery due to gut ischemia (narrowing or blockage of blood vessels). Such observations, although preliminary, lead to the question of what role human gut phages are playing in specific diseases, and in particular the microbial dysbiosis often associated with a range of diseases.
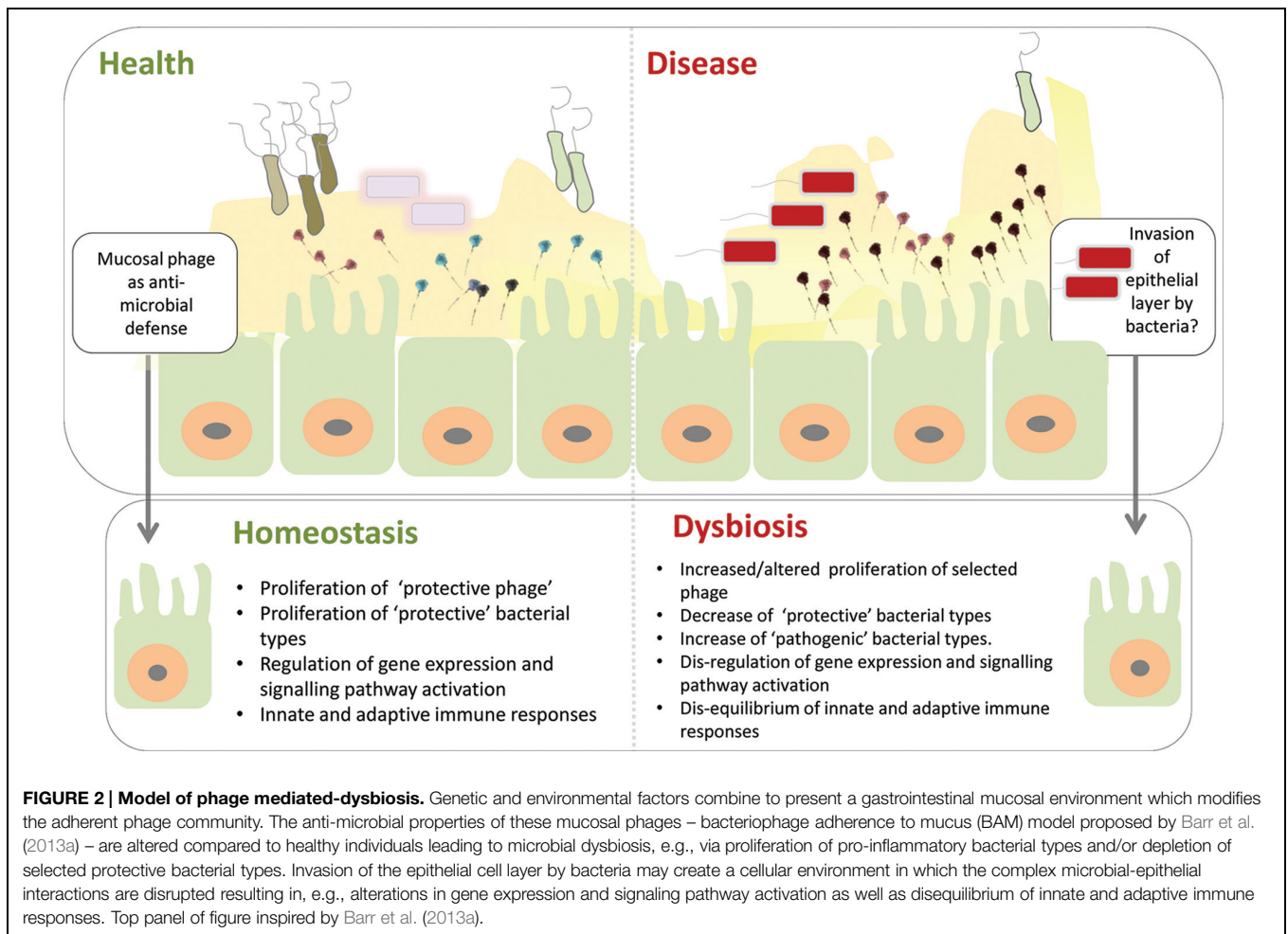
Barr et al. (2013a) recently posited a key role for phages in human health, suggesting that phages residing in metazoan mucosal surfaces provide a non-host derived anti-microbial defense. The model suggests that metazoan mucosal surfaces and phages coevolve to maintain phage adherence. This benefits the metazoan host by limiting mucosal bacteria, and benefits the phage through more frequent interactions with bacterial hosts. The study revealed that in divergent animal mucosa, phages were found to accumulate at significantly higher densities in comparison to non-mucosal surfaces. The observation that phage abundance is high in mucosal surfaces is of course not

ground-breaking in itself and could simply be explained through increased replication of phages due to the higher densities of bacterial hosts they encounter in these mucosal environments. Barr et al. (2013a), however, showed that even in the absence of host bacteria, phage abundance was greater on cells producing mucus in comparison to non-mucus producing cells. The authors also ruled out the likelihood that the accumulation of phage was simply the result of the laws of mass-action, i.e., the gel-like properties of the mucus itself slowing the diffusion of phage. Finally, the authors showed that enrichment of phage on mucosal surfaces was occurring via interactions between host mucin glycoproteins and phage immunoglobulin-like protein domains exposed on phage capsids. These observations suggest a symbiotic relationship between phages and metazoan hosts that provides a previously unrecognized antimicrobial defense that actively protects mucosal surfaces (Barr et al., 2013a). It must be noted, however, that the study by Barr et al. (2013a) did not involve human gut samples. Whether or not such a symbiotic relationship exists within the human gut has still to be shown. Moreover, other potential roles of phage within mucosal layers cannot be excluded, such as a survival strategy allowing phage to persist for long periods (in the absence of producing virions), until a suitable host is found (Edlund et al., 2015).

With these caveats in mind it is still tempting to speculate and hypothesize that disruption of this symbiosis may be a significant component of the microbial dysbiosis signature associated with a range of human diseases. In this scenario the human host genetic background and/or lifestyle factors may lead to an altered representation of anti-microbial mucosal phage, e.g., increased abundance of phages leading to selected reduction in bacterial species, potentially propelling the microbial community into dysbiosis, leading to deregulation of protective microbial-epithelial cell interactions (Hooper and Gordon, 2001; Kostic et al., 2013). Such deregulation of the complex and finely tuned interactions associated with intestinal immune responses could potentially allow invasion of 'dysbiotic' bacterial intruders through the epithelial cell layer (Hooper and Gordon, 2001; Kostic et al., 2013; see **Figure 2**). Realization that phages may contribute to a non-host derived immunity, provides a range of possibilities to modulate this mucosal phage complement through targeted interventions to prevent and or treat intestinal diseases (see also Barr et al., 2013b).

Future Outlook

With advances in high-throughput sequencing technologies and associated methods for sequence analysis and assignment, we are now starting to gain some crucial insight into the unique ecological characteristics of the human gut virome, shedding light on its structure and functioning. Although not without limitations, the application of these novel methodologies is providing the potential to gain unprecedented access to this human gut collective, allowing us to further explore the complex nature of the phage-host relationship. Concepts and principles that will aid our understanding of this environment are in the

FIGURE 2 | Model of phage mediated-dysbiosis. Genetic and environmental factors combine to present a gastrointestinal mucosal environment which modifies the adherent phage community. The anti-microbial properties of these mucosal phages – bacteriophage adherence to mucus (BAM) model proposed by Barr et al. (2013a) – are altered compared to healthy individuals leading to microbial dysbiosis, e.g., via proliferation of pro-inflammatory bacterial types and/or depletion of selected protective bacterial types. Invasion of the epithelial cell layer by bacteria may create a cellular environment in which the complex microbial-epithelial interactions are disrupted resulting in, e.g., alterations in gene expression and signaling pathway activation as well as disequilibrium of innate and adaptive immune responses. Top panel of figure inspired by Barr et al. (2013a).

process of being formed, shaped by incremental increases in knowledge that place gut phages on a continuum from simple gene transfer agents to community level functional buffers, and potential co-regulators of human immune defense systems and drivers of ecosystem diversity.

It is evident that viral metagenomics faces many major challenges, not least the issues of sequence analysis and classification, but current solutions and approaches to analyzing these datasets are widening our perspectives on human gut ecology at a rapid pace. We must, however, proceed with caution, and heed the lessons learned from microbial ecology over the last decade. Wholescale investment in metagenomic approaches, powerful as they are, at the expense of the established more traditional-culture based methods, is unlikely to deliver the full story and we must not progress in one methodology at the expense of skills and capacity in underpinning basic virology. A combination of approaches, including emerging approaches for studying phage-host interactions (Dang and Sullivan, 2014), is necessary for a complete understanding and, perhaps most importantly, for application of knowledge gained.

With new insights will also come more questions. The suggestion that gut mucosal phages could enter into a mutualistic relationship with their human host directly (without working through the indirect manipulation of community structure and output), essentially suggests the existence of a hitherto uncharacterized branch of the human immune system (Barr et al., 2013a). Further proof of such functionality may actually force us to further refine our understanding of the inherent ecological characteristics of the gut ecosystem (as well as our assessment of components comprising the human holobiont), and opens for manipulation an additional range of complex interactions between human host, microbiome, and virome that may be used to enhance human health and well-being. Without doubt, further understanding of the role of human gut-specific phages within healthy and diseased individuals is crucial to the successful harnessing of the full diagnostic and/or therapeutic potential of the human gut virome.

Acknowledgments

References

Abeles, S. R., Robles-Sikisaka, R., Ly, M., Lum, A. G., Salzman, J., Boehm, T. K., et al. (2014). Human oral viruses are personal, persistent and gender-consistent. *ISME J.* 8, 1753–1767. doi: 10.1038/ismej.2014.31

Ahmed, S., Macfarlane, G. T., Fite, A., McBain, A. J., Gilbert, P., and Macfarlane, S. (2007). Mucosa-associated bacterial diversity in relation to human terminal ileum and colonic biopsy samples. *Appl. Environ. Microbiol.* 73, 7435–7442. doi: 10.1128/AEM.01143-07

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Amann, R. I., Ludwig, W., and Schleifer, K. H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* 59, 143–169.

Angly, F. E., Felts, B., Breitbart, M., Salamon, P., Edwards, R. A., Carlson, C., et al. (2006). The marine viromes of four oceanic regions. *PLoS Biol.* 4:e368. doi: 10.1371/journal.pbio.0040368

Angly, F., Rodriguez-Brito, B., Bangor, D., McNairnie, P., Breitbart, M., Salamon, P., et al. (2005). PHACCS, an online tool for estimating the structure and diversity of uncultured viral communities using metagenomic information. *BMC Bioinform.* 6:41. doi: 10.1186/1471-2105-6-41

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D. R., et al. (2011). Enterotypes of the human gut microbiome. *Nature* 473, 174–180. doi: 10.1038/nature09944

Bäckhed, F., Ley, R. E., Sonnenburg, J. L., Peterson, D. A., and Gordon, J. I. (2005). Host-bacterial mutualism in the human intestine. *Science* 307, 1915–1920. doi: 10.1126/science.1104816

Barr, J. J., Auro, R., Furlan, M., Whiteson, K. L., Erb, M. L., Pogliano, J., et al. (2013a). Bacteriophage adhering to mucus provide a non-host-derived immunity. *Proc. Natl. Acad. Sci. U.S.A.* 110, 10771–10776. doi: 10.1073/pnas.1305923110

Barr, J. J., Youle, M., and Rohwer, F. (2013b). Innate and acquired bacteriophage-mediated immunity. *Bacteriophage* 3:e25857. doi: 10.4161/bact.25857

Bibby, K. (2014). Improved bacteriophage genome data is necessary for integrating viral and bacterial ecology. *Microb. Ecol.* 67, 242–244. doi: 10.1007/s00248-013-0325-x

Breitbart, M. (2012). Marine viruses, truth or dare. *Annu. Rev. Mar. Sci.* 4, 425–448. doi: 10.1146/annurev-marine-120709-142805

Breitbart, M., Haynes, M., Kelley, S., Angly, F., Edwards, R. A., Felts, B., et al. (2008). Viral diversity and dynamics in an infant gut. *Res. Microbiol.* 159, 367–373. doi: 10.1016/j.resmic.2008.04.006

Breitbart, M., Hewson, I., Felts, B., Mahaffy, J. M., Nulton, J., Salamon, P., et al. (2003). Metagenomic analyses of an uncultured viral community from human feces. *J. Bacteriol.* 85, 6220–6223. doi: 10.1128/JB.185.20.6220-6223.2003

Breitbart, M., Miyake, J. H., and Rohwer, F. (2004). Global distribution of nearly identical phage-encoded DNA sequences. *FEMS Microbiol. Lett.* 236, 249–256. doi: 10.1111/j.1574-6968.2004.tb09654.x

Breitbart, M., and Rohwer, F. (2005). Here a virus, there a virus, everywhere the same virus? *Trends Microbiol.* 13, 278–284. doi: 10.1016/j.tim.2005.04.003

Brüssow, H. (2013). Bacteriophage-host interaction, from splendid isolation into a messy reality. *Curr. Opin. Microbiol.* 16, 500–506. doi: 10.1016/j.mib.2013.04.007

Brüssow, H., Canchaya, C., and Hardt, W.-D. (2004). Phages and the evolution of bacterial pathogens, from genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev.* 68, 560–602. doi: 10.1128/MMBR.68.3.560-602.2004

Brüssow, H., and Hendrix, R. W. (2002). Phage genomics, small is beautiful. *Cell* 108, 13–16. doi: 10.1016/S0092-8674(01)00637-7

Casjens, S. (2003). Prophages and bacterial genomics, what have we learned so far? *Mol. Microbiol.* 49, 277–300. doi: 10.1046/j.1365-2958.2003.03580.x

Chibani-Chennoufi, S., Bruttin, A., Dillmann, M.-L., and Brüssow, H. (2004). Phage-host interaction, an ecological perspective. *J. Bacteriol.* 186, 3677–3686. doi: 10.1128/JB.186.12.3677-3686.2004

Clemente, J. C., Ursell, L. K., Parfrey, L. W., and Knight, R. (2012). The impact of the gut microbiota on human health, an integrative view. *Cell* 148, 1258–1270. doi: 10.1016/j.cell.2012.01.035

Colomer-Lluch, M., Jofre, J., and Muniesa, M. (2011). Antibiotic resistance genes in the bacteriophage DNA fraction of environmental samples. *PLoS ONE* 6:e17549. doi: 10.1371/journal.pone.0017549

Cryan, J. F., and Dinan, T. G. (2012). Mind-altering microorganisms, the impact of the gut microbiota on brain and behaviour. *Nat. Rev. Neurosci.* 13, 701–712. doi: 10.1038/nrn3346

Dang, V. T., and Sullivan, M. B. (2014). Emerging methods to study bacteriophage infection at the single cell level. *Front. Microbiol.* 5:724. doi: 10.3389/fmicb.2014.00724

d'Herelle, F. (1917). Sur un microbe invisible antagoniste des bacilles dysentérique. *C.R. Acad. Sci. Paris* 165, 373–375.

Dinakaran, V., Rathinavel, A., Pushpanathan, M., Sivakumar, R., Gunasekaran, P., and Rajendhran, J. (2014). Elevated levels of circulating DNA in cardiovascular disease patients, metagenomic profiling of microbiome in the circulation. *PLoS ONE* 9:e105221. doi: 10.1371/journal.pone.0105221

Duerkop, B. A., Clements, C. V., Rollins, D., Rodrigues, J. L. M., and Hooper, L. V. (2012). A composite bacteriophage alters colonization by an intestinal commensal bacterium. *Proc. Natl. Acad. Sci. U.S.A.* 109, 17621–11766. doi: 10.1073/pnas.1206136109

Dutilh, B. E., Cassman, N., McNair, K., Sanchez, S. E., Silva, G. G. Z., Boling, L., et al. (2014). A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nat. Commun.* 5, 4498. doi: 10.1038/ncomms5498

Dutilh, B. E., Schmieder, R., Nulton, J., Felts, B., Salamon, P., Edwards, R. A., et al. (2012). Reference-independent comparative metagenomics using cross-assembly, crAss. *Bioinformatics* 28, 3225–3231. doi: 10.1093/bioinformatics/bts613

Edlund, A., Santiago-Rodriguez, T. M., Boehm, T. K., and Pride, D. T. (2015). Bacteriophage and their potential roles in the human oral cavity. *J. Oral. Microbiol.* 7, 27423. doi: 10.3402/jom.v7.27423

Flewett, T. H., Bryden, A. S., and Davies, H. (1974). Diagnostic electron microscopy of faeces. I. The viral flora of the faeces as seen by electron microscopy. *J. Clin. Pathol.* 27, 603–608. doi: 10.1136/jcp.27.8.603

Fuhrman, J. A. (1999). Marine viruses and their biogeochemical and ecological effects. *Nature* 399, 541–548. doi: 10.1038/21119

Furuse, K., Osawa, S., Kawashiro, J., Tanaka, R., Ozawa, A., Sawamura, S., et al. (1983). Bacteriophage distribution in human faeces, continuous survey of healthy subjects and patients with internal and leukaemic diseases. *J. Gen. Virol.* 64, 2039–2043. doi: 10.1099/0022-1317-64-9-2039

Gorski, A., Dabrowska, K., Switala-Jeleń, K., Nowaczyk, M., Weber-Dabrowska, B., Boratynski, J., et al. (2003). New insights into the possible role of bacteriophages in host defense and disease. *Med. Immunol.* 2, 2. doi: 10.1186/1476-9433-2-2

Hawkins, S. A., Layton, A. C., Ripp, S., Williams, D., and Sayler, G. S. (2008). Genome sequence of the *Bacteroides fragilis* phage ATCC 51477-B1. *Virol. J.* 5, 97. doi: 10.1186/1743-422X-5-97

Hooper, L. V., and Gordon, J. I. (2001). Commensal host-bacterial relationships in the gut. *Science* 292, 1115–1118. doi: 10.1126/science.1058709

Hooper, L. V., and Macpherson, A. J. (2010). Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nat. Rev. Immunol.* 10, 159–169. doi: 10.1038/nri2710

Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J. H., Chinwalla, A. T., et al. (2012). Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214. doi: 10.1038/nature11234

Jones, B. V. (2010). The human gut mobile metagenome, a metazoan perspective. *Gut Microbes* 1, 415–431. doi: 10.4161/gmic.1.6.14087

Jones, B. V., and Marchesi, J. R. (2007). Accessing the mobile metagenome of the human gut microbiota. *Mol. Biosyst.* 3, 749–758. doi: 10.1039/b705657e

Kim, M.-S., Park, E.-J., Roh, S. W., and Bae, J.-W. (2011). Diversity and abundance of single-stranded DNA viruses in human feces. *Appl. Environ. Microbiol.* 77, 8062–8070. doi: 10.1128/AEM.06331-11

Koskella, B., and Brockhurst, M. A. (2014). Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol. Rev.* 38, 916–931. doi: 10.1111/1574-6976.12072

Kostic, A. D., Howitt, M. R., and Garrett, W. S. (2013). Exploring host-microbiota interactions in animal models and humans. *Genes Dev.* 27, 701–718. doi: 10.1101/gad.212522.112

Lepage, P., Colombet, J., Marteau, P., Sime-Ngando, T., Doré, J., and Leclerc, M. (2008). Dysbiosis in inflammatory bowel disease, a role for bacteriophages? *Gut* 57, 424–425. doi: 10.1136/gut.2007.134668

Letarov, A., and Kulikov, E. (2009). The bacteriophages in human- and animal body-associated microbial communities. *J. Appl. Microbiol.* 107, 1–13. doi: 10.1111/j.1365-2672.2009.04143.x

Ley, R. E., Peterson, D. A., and Gordon, J. I. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* 124, 837–848. doi: 10.1016/j.cell.2006.02.017

Licht, T. R., Christensen, B. B., Krogfelt, K. A., and Molin, S. (1999). Plasmid transfer in the animal intestine and other dynamic bacterial populations, the role of community structure and environment. *Microbiology* 145, 2615–2622.

Licht, T. R., and Wilcks, A. (2005). Conjugative gene transfer in the gastrointestinal environment. *Adv. Appl. Microbiol.* 58C, 77–95. doi: 10.1016/S0065-2164(05)58002-X

Looft, T., Allen, H. K., Cantarel, B. L., Levine, U. Y., Bayles, D. O., Alt, D. P., et al. (2014). Bacteria, phages and pigs, the effects of in-feed antibiotics on the microbiome at different gut locations. *ISME J.* 8, 1566–1576. doi: 10.1038/ismej.2014.12

Marti, E., Variatza, E., and Balcázar, J. L. (2014). Bacteriophages as a reservoir of extended-spectrum β-lactamase and fluoroquinolone resistance genes in the environment. *Clin. Microbiol. Infect. Dis.* 20, O456–O459. doi: 10.1111/1469-0691.12446

Mazmanian, S. K., Round, J. L., and Kasper, D. L. (2008). A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature* 453, 620–625. doi: 10.1038/nature07008

Mello, L. V., Chen, X., and Rigden, D. J. (2010). Mining metagenomic data for novel domains, BACON, a new carbohydrate-binding module. *FEBS Lett.* 584, 2421–2426. doi: 10.1016/j.febslet.2010.04.045

Minot, S., Bryson, A., Chehoud, C., Wu, G. D., Lewis, J. D., and Bushman, F. D. (2013). Rapid evolution of the human gut virome. *Proc. Natl. Acad. Sci. U.S.A.* 110, 12450–12455. doi: 10.1073/pnas.1300833110

Minot, S., Sinha, R., Chen, J., Li, H., Keilbaugh, S. A., Wu, G. D., et al. (2011). The human gut virome, inter-individual variation and dynamic response to diet. *Genome Res.* 21, 1616–1625. doi: 10.1101/gr.122705.111

Minot, S., Wu, G. D., Lewis, J. D., and Bushman, F. D. (2012a). Conservation of gene cassettes among diverse viruses of the human gut. *PLoS ONE* 7:e42342. doi: 10.1371/journal.pone.0042342

Minot, S., Grunberg, S., Wu, G. D., Lewis, J. D., and Bushman, F. D. (2012b). Hypervariable loci in the human gut virome. *Proc. Natl. Acad. Sci. U.S.A.* 109, 3962–3966. doi: 10.1073/pnas.1119061109

Modi, S. R., Lee, H. H., Spina, C. S., and Collins, J. J. (2013). Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. *Nature* 499, 219–222. doi: 10.1038/nature12212

Muniesa, M., and Jofre, J. (2014). Identifying and analyzing bacteriophages in human fecal samples, what could we discover? *Future Microbiol.* 9, 879–886. doi: 10.2217/fmb.14.47

Norman, J. M., Handley, S. A., Baldridge, M. T., Droit, L., Liu, C. Y., Keller, B. C., et al. (2015). Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell* 160, 447–460. doi: 10.1016/j.cell.2015.01.002

Ogilvie, L. A., Bowler, L. D., Caplin, J., Dedi, C., Diston, D., Cheek, E., et al. (2013). Genome signature-based dissection of human gut metagenomes to extract subliminal viral sequences. *Nat. Commun.* 4, 2420. doi: 10.1038/ncomms3420

Ogilvie, L. A., Caplin, J., Dedi, C., Diston, D., Cheek, E., Bowler, L., et al. (2012a). Comparative (meta) genomic analysis and ecological profiling of human gut-specific bacteriophage φB124-14. *PLoS ONE* 7:e35053. doi: 10.1371/journal.pone.0035053

Ogilvie, L. A., Firouzmand, S., and Jones, B. V. (2012b). Evolutionary, ecological and biotechnological perspectives on plasmids resident in the human gut mobile metagenome. *Bioeng. Bugs* 3, 13–31. doi: 10.4161/bbug.3.1.17883

Ogilvie, L. A., and Hirsch, P. (2012). *Microbial Ecological Theory: Current Perspectives.* Norwich: Caister Academic Press, 120.

Ogilvie, L. A., and Jones, B. V. (2013). "The human gut mobile metagenome, a metazoan perspective," in *Encyclopedia of Metagenomics,* ed. S. Highlander (New York: Springer Science & Business Media), 1–14.

O'Hara, A. M., and Shanahan, F. (2006). The gut flora as a forgotten organ. *EMBO Rep.* 7, 688–693. doi: 10.1038/sj.embor.7400731

Ohman, L., Törnblom, H., and Simrén, M. (2015). Crosstalk at the mucosal border, importance of the gut microenvironment in IBS. *Nat. Rev. Gastroenterol. Hepatol.* 12, 36–49. doi: 10.1038/nrgastro.2014.200

Paterson, S., Vogwill, T., Buckling, A., Benmayor, R., Spiers, A. J., Thomson, N. R., et al. (2010). Antagonistic coevolution accelerates molecular evolution. *Nature* 464, 275–278. doi: 10.1038/nature08798

Pérez-Brocal, V., García-López, R., Vázquez-Castellanos, J. F., Nos, P., Beltrán, B., Latorre, A., et al. (2013). Study of the viral and microbial communities associated with Crohn's disease, a metagenomic approach. *Clin. Transl. Gastroenterol.* 4:e36. doi: 10.1038/ctg.2013.9

Pride, D. T., Meinersmann, R. J., Wassenaar, T. M., and Blaser, M. J. (2003). Evolutionary implications of microbial genome tetranucleotide frequency biases. *Genome Res.* 13, 145–158. doi: 10.1101/gr.335003

Pride, D. T., Salzman, J., Haynes, M., Rohwer, F., Davis-Long, C., White, R. A. III, et al. (2012). Evidence of a robust resident bacteriophage population revealed through analysis of the human salivary virome. *ISME J.* 6, 915–926. doi: 10.1038/ismej.2011.169

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65. doi: 10.1038/nature08821

Rajilić-Stojanović, M., and de Vos, W. M. (2014). The first 1000 cultured species of the human gastrointestinal microbiota. *FEMS Microbiol. Rev.* 38, 996–1047. doi: 10.1111/1574-6976.12075

Rakoff-Nahoum, S., Paglino, J., Eslami-Varzaneh, F., Edberg, S., and Medzhitov, R. (2004). Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. *Cell* 118, 229–241. doi: 10.1016/j.cell.2004.07.002

Rappé, M. S., and Giovannoni, S. J. (2003). The uncultured microbial majority. *Annu. Rev. Microbiol.* 57, 369–394. doi: 10.1146/annurev.micro.57.030502.090759

Reyes, A., Haynes, M., Hanson, N., Angly, F. E., Heath, A. C., Rohwer, F., et al. (2010). Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature* 466, 334–338. doi: 10.1038/nature09199

Reyes, A., Semenkovich, N. P., Whiteson, K., Rohwer, F., and Gordon, J. I. (2012). Going viral, next-generation sequencing applied to phage populations in the human gut. *Nat. Rev. Microbiol.* 10, 607–617. doi: 10.1038/nrmicro2853

Reyes, A., Wu, M., McNulty, N. P., Rohwer, F. L., and Gordon, J. I. (2013). Gnotobiotic mouse model of phage-bacterial host dynamics in the human gut. *Proc. Natl. Acad. Sci. U.S.A.* 110, 20236–20241. doi: 10.1073/pnas.1319470110

Rohwer, F., and Thurber, R. V. (2009). Viruses manipulate the marine environment. *Nature* 459, 207–212. doi: 10.1038/nature08060

Roux, S., Tournayre, J., Mahul, A., Debroas, D., and Enault, F. (2014). Metavir 2, new tools for viral metagenome comparison and assembled virome analysis. *BMC Bioinform.* 15:76. doi: 10.1186/1471-2105-15-76

Sano, E., Carlson, S., Wegley, L., and Rohwer, F. (2004). Movement of viruses between biomes. *Appl. Environ. Microbiol.* 70, 5842–5846. doi: 10.1128/AEM.70.10.5842-5846.2004

Seguritan, V., Alves, N., Arnoult, M., Raymond, A., Lorimer, D., Burgin, A. B., et al. (2012). Artificial neural networks trained to detect viral and phage structural proteins. *PLoS Comput. Biol.* 8:e1002657. doi: 10.1371/journal.pcbi.1002657

Smillie, C. S., Smith, M. B., Friedman, J., Cordero, O. X., David, L. A., and Alm, E. J. (2011). Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* 480, 241–244. doi: 10.1038/nature10571

Sobhani, I., Tap, J., Roudot-Thoraval, F., Roperch, J. P., Letulle, S., Langella, P., et al. (2011). Microbial dysbiosis in colorectal cancer (CRC) patients. *PLoS ONE* 6:e16393. doi: 10.1371/journal.pone.0016393

Sokol, H., Pigneur, B., Watterlot, L., Lakhdari, O., Bermúdez-Humarán, L. G., Gratadoux, J.-J., et al. (2008). *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc. Natl. Acad. Sci. U.S.A.* 105, 16731–16736. doi: 10.1073/pnas.0804812105

Sorek, R., Kunin, V., and Hugenholtz, P. (2008). CRISPR–a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat. Rev. Microbiol.* 6, 181–186. doi: 10.1038/nrmicro1793

Stern, A., Mick, E., Tirosh, I., Sagy, O., and Sorek, R. (2012). CRISPR targeting reveals a reservoir of common phages associated with the human gut microbiome. *Genome Res.* 22, 1985–1994. doi: 10.1101/gr.138297.112

Suau, A., Bonnet, R., Sutren, M., Godon, J. J., Gibson, G. R., Collins, M. D., et al. (1999). Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. *Appl. Environ. Microbiol.* 65, 4799–4807.

Suttle, C. A. (2005a). Viruses in the sea. *Nature* 437, 356–361. doi: 10.1038/nature04160

Suttle, C. A. (2005b). Crystal ball. The viriosphere, the greatest biological diversity on Earth and driver of global processes. *Environ. Microbiol.* 7, 481–482. doi: 10.1111/j.1462-2920.2005.803_11.x

Suttle, C. A. (2007). Marine viruses–major players in the global ecosystem. *Nat. Rev. Microbiol.* 5, 801–812. doi: 10.1038/nrmicro1750

Teeling, H., Waldmann, J., Lombardot, T., Bauer, M., and Glöckner, F. O. (2004). TETRA, a web-service and a stand-alone program for the analysis and comparison of tetranucleotide usage patterns in DNA sequences. *BMC Bioinform.* 5:163. doi: 10.1186/1471-2105-5-163

Thingstad, T. F., Bratbak, G., and Heldal, M. (2008). "Aquatic Phage Ecology," in *Bacteriophage Ecology, Population Growth, Evolution, and Impact of Bacterial Viruses*, ed. S. T. Abedon (Cambridge: Cambridge University Press), 251–280. doi: 10.1017/CBO9780511541483.013

Thurber, R. V., Haynes, M., Breitbart, M., Wegley, L., and Rohwer, F. (2009). Laboratory procedures to generate viral metagenomes. *Nat. Protoc.* 4, 470–483. doi: 10.1038/nprot.2009.10

Twort, F. W. (1915). An investigation on the nature of ultra-microscopic viruses. *Lancet* 186, 1241–1243. doi: 10.1016/S0140-6736(01)20383-3

Wagner, J., Maksimovic, J., Farries, G., Sim, W. H., Bishop, R. F., Cameron, D. J., et al. (2013). Bacteriophages in gut samples from pediatric Crohn's disease patients, metagenomic analysis using 454 pyrosequencing. *Inflamm. Bowel. Dis.* 19, 1598–1608.

Wang, T., Cai, G., Qiu, Y., Fei, N., Zhang, M., Pang, X., et al. (2012). Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. *ISME J.* 6, 320–329. doi: 10.1038/ismej.2011.109

Wang, X., Kim, Y., Ma, Q., Hong, S. H., Pokusaeva, K., Sturino, J. M., et al. (2010). Cryptic prophages help bacteria cope with adverse environments. *Nat. Commun.* 1, 147. doi: 10.1038/ncomms1146

Weinbauer, M. G. (2004). Ecology of prokaryotic viruses. *FEMS Microbiol. Rev.* 28, 127–181. doi: 10.1016/j.femsre.2003.08.001

Wexler, H. M. (2007). Bacteroides, the good, the bad, and the nitty-gritty. *Clin. Microbiol. Rev.* 20, 593–621. doi: 10.1128/CMR.00008-07

Whitman, W. B., Coleman, D. C., and Wiebe, W. J. (1998). Prokaryotes, the unseen majority. *Proc. Natl. Acad. Sci. U.S.A.* 95, 6578–6583. doi: 10.1073/pnas.95.12.6578

Wommack, K. E., Bhavsar, J., Polson, S. W., Chen, J., Dumas, M., Srinivasiah, S., et al. (2012). VIROME, a standard operating procedure for analysis of viral metagenome sequences. *Stand. Genomic Sci.* 6, 427–439. doi: 10.4056/sigs.2945050

Wommack, K. E., and Colwell, R. R. (2000). Virioplankton, viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* 64, 69–114. doi: 10.1128/MMBR.64.1.69-114.2000

Wommack, K. E., Ravel, J., Hill, R. T., and Colwell, R. R. (1999). Hybridization analysis of chesapeake bay virioplankton. *Appl. Environ. Microbiol.* 65, 241–250.

Wu, G. D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.-Y., Keilbaugh, S. A., et al. (2011). Linking long-term dietary patterns with gut microbial enterotypes. *Science* 334, 105–108. doi: 10.1126/science.1208344

Zhang, T., Breitbart, M., Lee, W. H., Run, J.-Q., Wei, C. L., Soh, S. W., et al. (2006). RNA viral community in human feces: prevalence of plant pathogenic viruses. *PLoS Biol.* 4:e3. doi: 10.1371/journal.pbio.0040003

Zhang, Z., Geng, J., Tang, X., Fan, H., Xu, J., Wen, X., et al. (2014). Spatial heterogeneity and co-occurrence patterns of human mucosal-associated intestinal microbiota. *ISME J.* 8, 881–893. doi: 10.1038/ismej.2013.185

Zoetendal, E. G., Vaughan, E. E., and de Vos, W. M. (2006). A microbial world within us. *Mol. Microbiol.* 59, 1639–1650.