The healthy human virome: from virus–host symbiosis to disease

Eugene V Koonin, Valerian Dolja, Mart Krupovic

The healthy human virome: from virus-host symbiosis to disease

Eugene V. Koonin,[1,#] Valerian V. Dolja,[2] Mart Krupovic[3]

[1]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland, USA

[2]Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon, USA

[3]Archaeal Virology Unit, Institut Pasteur, Paris, France

[#]For correspondence: koonin@ncbi.nlm.nih.gov

Abstract

Viruses are ubiquitous, essential components of any ecosystem, and of multicellular organism holobionts. Numerous viruses cause acute infection, killing the host or being cleared by immune system. In many other cases, viruses coexist with the host as symbionts, either temporarily or for the duration of the host's life. Apparently, virus-host relationships span the entire range from aggressive parasitism to mutualism. Here we attempt to delineate the healthy human virome, that is, the entirety of viruses that are present in a healthy human body. The bulk of the healthy virome consists of bacteriophages infecting bacteria in the intestine and other locations. However, a variety of viruses, such as anelloviruses and herpesviruses, and the numerous endogenous retroviruses, persist by replicating in human cells, and these are our primary focus. Crucially, the boundary between symbiotic and pathogenic viruses is fluid such that members of the healthy virome can become pathogens under changing conditions.

Introduction

Billions of years of perennial and ubiquitous co-evolution between viruses and cells have produced a broad spectrum of virus-host interaction regimes, ranging from aggressive antagonism to commensalism, whereby viruses coexist with their hosts without harming them, at least, in the short term, and even to mutualism when a virus is beneficial and can be essential to the host [1-4]. The relationship between a particular virus and its host can be rarely, if ever, defined by a single regime. Rather, the mode of virus-host interaction is a function of multiple factors, including the environmental conditions, host and virus population structure, the immunological status of the host and many more.

Virus-host relationships that do not result in the demise of the host cell have been described across the virosphere [3,4]. Hence the concept of the 'normal' or 'healthy' virome, in principle, is applicable to any organism, including humans. Indeed, although most of the best characterized human viruses cause acute infections and are associated with diseases, a healthy human organism is host to a much greater variety of viruses [5-7]. The substantial majority of these infect bacteria that inhabit the human intestine, but a number of viruses actually reproduce in human cells without causing disease, at least, in the short term.

In this brief review, we systematically discuss the healthy human virome and emphasize that the boundary between "normal" and pathogenic (that is, causally associated with clinically manifested disease) viromes is blurred. Indeed, the same virus can be either a symbiont (with either no perceptible fitness effect on the host, that is, effectively, a commensal, or with a beneficial effect) or a pathogen depending on the conditions such as the health status and developmental stage of the host. We further stress that the current knowledge of symbiotic viruses lags far behind that of the pathogenic viruses.

The healthy human phageome

The human intestine, by far the richest microbial habitat in the body, contains about 10^{14} bacterial and archaeal cells at any given moment [8]. As many microbial communities, the human microbiome hosts a broad variety of viruses, in which tailed bacteriophages (class *Caudoviricetes* within the realm *Duplodnaviria*) comprise the overwhelming majority [8,9], albeit with a considerable contribution by ssDNA phages of the families *Microviridae* [10-12] and *Inoviridae* [13] (realm *Monodnaviria*). Systematic metagenomics surveys of the human phageome identify

thousands of phages but show that relatively few are common components of the healthy gut phageome. Thus, one of the most detailed metagenomics analyses resulted in the identification of only 23 phages that were shared by >50% of the tested individuals [9]. Strikingly, the most prevalent human-associated phage (and most prominent component of the healthy human virome altogether), the crAssphage, has been discovered only through metagenomics [14]. This initial finding has been followed by the characterization of an expansive group of crAss-like phages, also by metagenome analysis [15,16]. Subsequently, some crAss-like phages have been grown in cultures of the respective host bacteria, members of the phylum Bacteroidetes [17,18]. Given the difficulty of growing the human intestine-associated bacteria-phage systems in the laboratory, the actual size and diversity of the healthy human phageome remains to be discovered. Importantly, substantial changes in the human phageome have been associated with various diseases conditions, including infection with human and simian immunodeficiency viruses, and in some cases, the phageome perturbation was, apparently, decoupled from the changes in the microbiome [19-21].

Symbiotic viruses replicating in human cells

Over many decades, diverse viruses have been isolated from healthy humans more or less by chance. In the last few years, systematic surveys of the healthy human virome became possible thanks to the advances of metagenomics [6,7]. Below we briefly discuss the viruses that have been shown, either by traditional virology approaches or by metagenomics (Figure 1a), to be commonly present in healthy humans without causing apparent disease, following the main divisions of the current virus taxonomy [22,23]. The key aspects of the association of these viruses with the human organism are summarized in Figure 1.

Realm *Riboviria*

Kingdom *Orthornavira*

The realm *Riboviria* consists of viruses with RNA genomes as well as viruses with DNA genomes that employ reverse transcription in their replication cycles. The kingdom *Orthornavira* encompasses RNA viruses that share homologous RNA-dependent RNA polymerases. Many RNA viruses have been isolated or detected by metatranscriptomics in healthy humans but few can be confidently identified as symbionts replicating in human cells (Figure 1a).

Perhaps, the most notable apparent human symbionts among the orthornaviruses are pegiviruses (family *Flaviviridae*) [24,25]. The prototypical pegivirus has been initially tentatively identified

as hepatitis G virus [26], but no association with hepatitis has been subsequently confirmed [27]. In the family tree of flavivirids, pegiviruses tightly cluster with hepaciviruses which include hepatitis C virus (HCV), a major human pathogen [24,25]. However, unlike HCV, pegiviruses have not been linked to any pathology. Pegivirus infection in humans is common, with the incidence of about 5%, and pegiviruses readily grow in human cell cultures, leaving no doubt that these are bona fide human viruses [25]. Notably, pegivirus infection appears to be associated with benign clinical outcome in AIDS patients [28,29] indicating that these apparent viral symbionts of humans could benefit the host via protection from other viruses.

A recent, intriguing addition to the human virome are statoviruses (for STool-Associated TOmbus-like viruses) that were originally identified in multiple metagenomes from humans, macaques, cows and mice [30], followed by detection in nasal-throat swabs of humans with acute respiratory disease [31]. Although the actual hosts of statoviruses remain unknown, RdRP phylogeny, where statovirus RdRPs cluster with a variety of unclassified tombus-like viruses from invertebrate holobionts (that is, a host together with the entirety of associated symbionts and parasites) [32], suggests that statoviruses are associated with either mammalian diet (e.g., plants) or, more likely, some protist symbionts or parasites.

Another widespread group of putative members of the healthy human virome are members of *Picobirnaviridae*. Picobirnaviruses are commonly detected in mammalian, including human, intestines and have not been convincingly linked to any diseases although are suspected to be associated with diarrhea [33,34]. Picobirnaviruses have never been grown in cell cultures, and their true hosts remain unknown. The presence of highly conserved ribosome-binding sites (Shine-Dalgarno sequences), which are a hallmark of prokaryotic mRNAs, has led to the suggestion that picobirnaviruses infect bacteria in mammalian microbiomes [35].

In addition, certain food-derived plant viruses are present in the human gastrointestinal tract and in feces, sometimes in substantial amounts. The list of such viruses is topped by pepper mild mottle virus (PMMoV), a tobamovirus commonly found in pepper, pepper-derived products and their consumers all over the world [36]. Remarkably, PMMoV retains infectivity after passing through the human alimentary tract. Due to its stability and wide presence in human feces, PMMoV is used as a surrogate marker of fecal contamination of water [37]. However, given that there is no

evidence of plant virus replication in vertebrate cells, any substantial role of plant viruses in human health is an extremely remote possibility.

Kingdom *Pararnavira*

This virus kingdom includes viruses that employ reverse transcription in their replication cycles. Over 3,000 of human endogenous retroviruses (HERVs) are integrated into the host genome, comprising about 8% of human DNA [38]. Accordingly, the HERVs play important and diverse roles in human biology that cannot be discussed in this brief review in any detail (for recent reviews, see [39-42]). Most of the HERVs appear to descend from ancient integration events so that the virus genes are disrupted and rearranged. The legacy of these ancient HERVs are genes encoding virus structural proteins (Gag and Env) that have been recruited for a variety of physiological functions [43], the best known being syncitins, the placental trophoblast receptors [39]. However, some members of the youngest group of HERVs, known as HERV-K or HML2, that are thought to have invaded the human genome less than a million years ago form virus particles, particularly, in early embryogenesis [44,45]. The HERV-K viruses, at least, are bona fide members of the healthy human virome. Many of the other HERVs are expressed as well and are implicated in a variety of functions including modulation of innate immunity, even though functional virus proteins are usually not produced [46]. In particular, it has been reported that expression of one of HERV-K suppresses the spread of invasive melanoma [47]. However, potential associations between HERVs and various diseases have been reported as well. Thus, the relationship between the HERVs and the human host is a typical symbiosis, with both beneficial and potential deleterious effects on the host (Figure 1b).

Realm *Monodnaviria*

Realm *Monodnaviria* includes prokaryotic and eukaryotic viruses which encode homologous replication initiation endonucleases of the HUH superfamily or their inactivated derivatives, as in the case of polyomaviruses and papillomaviruses [22,48]. In addition to phages of the *Microviridae* and *Inoviridae* families mentioned above, several other representatives of *Monodnaviria* have been repeatedly identified as part of the healthy human virome. These include members of the families *Parvoviridae*, *Genomoviridae*, *Smacoviridae*, *Papillomaviridae* and *Polyomaviridae* (Figure 1a). The actual hosts for human-associated genomoviruses [49] and smacoviruses [50] remain unknown but these viruses likely infect human-associated microbes rather than humans directly.

Parvoviruses

Parvoviruses from several genera have been detected in various samples from apparently healthy humans, including human bocaviruses (HBoV; genus *Bocaparvovirus*), adeno-associated viruses (AAVs; genus *Dependoparvovirus*), human parvovirus 4 (PARV4; genus *Tetraparvovirus*), parvovirus B19 (B19V; genus *Erythroparvovirus*) and several protoparvoviruses (genus *Protoparvovirus*) [51-54]. These viruses display highly variable cell tropism and pathogenicity. For instance, AAVs and PARV4 can infect cells from multiple tissue types and are not known to cause any disease. By contrast, HBoV is most commonly found in the respiratory and gastrointestinal tracts as well as in blood [55,56], and is associated with acute respiratory symptoms, especially in children [57]. However, the direct role of HBoV as a pathogen remains unclear. B19V invades red blood cell precursors in the bone marrow and is commonly considered as a human pathogen causing a wide range of pathological conditions, including the fifth disease in children, persistent anemia in immunocompromised patients, transient aplastic crises, hydrops fetalis in pregnant women, and arthropathy [52]. Yet, in healthy adults, B19V infections are largely asymptomatic [58], with the prevalence of up to 25% in healthy human skin biopsies [59]. Thus, B19V is a conditional component of the "healthy" human virome that can turn into a pathogen in response to various factors (Figure 1b).

Polyomaviruses and Papillomaviruses

Monodnaviria includes class *Papovaviricetes* that consists of *Polyomaviridae* and *Papillomaviridae*, two families of viruses with small (5-8 Kbps), circular dsDNA genomes [60,61]. Papovaviruses are thought to have evolved from parvoviruses [48], with polyomaviruses emerging in invertebrates and co-evolving with animals for at least half a billion years [62]. In humans, polyomaviruses lead a low-profile life styles characterized by low propagation levels, evasion of clearance by the immune system and asymptomatic infections in immunocompetent individuals [60]. Apparently, human polyomaviruses have evolved mechanisms to limit their own reproduction levels in order to establish persistent infections [63]. A poster child human polyomavirus is John Cunningham virus (JCV) that establishes life-long latent infections [64]. The seroprevalence of JCV and other human polyomaviruses in adults can reach 90% or higher, with common coinfection by several polyomaviruses that are transmitted through direct contacts between humans or through contaminated objects. Contrary to their name and the oncogenic potential of the large T (tumor) antigen (the early virus protein involved in genome replication) in

experimental settings, most of the human polyomaviruses are not oncogenic. The Merkel cell polyomavirus is the only one that has been convincingly identified as the etiological agent of the eponymous carcinoma, a skin cancer caused by malignant transformation of the skin neuroendocrine cells apparently facilitated by virus genome integration into the host chromosomes [60,65].

Papillomaviruses are the closest, albeit apparently somewhat 'younger' relatives of polyomaviruses that likely emerged in vertebrates ~350 mya, but their life style is distinct. The majority of the ~200 known human papillomaviruses (HPVs) belong to Beta and Gamma types (*Betapapillomavirus* and *Gammapapillomavirus* genera, respectively), and cause unapparent productive infections or low-grade disease of the skin or mucosal epithelium (e.g., warts and condylomas) that are normally cleared by the immune system [66,67]. The virus replication cycle is tightly linked to epithelial differentiation and is orchestrated by a regulatory network that involves coordination between the cell cycle, virus DNA replication and transcription, and RNA splicing [67,68]. Most of the infections by the 'high-risk', Alpha HPV types (HPV16 and HPV18 being most prevalent; genus *Alphapapillomavirus*) in women result in more prolonged cervical infections, 80-90% of which are asymptomatic and are eventually cleared by the immune system [66,67]. However, a small fraction of such infections, primarily due to defects in the immune response, progress to the formation of persistent papillomas and, in a minority of cases, to cervical cancer. This small fraction of HPV infections, nevertheless, accounts for close to 100% of cervical cancers that affect over 500,000 women annually [69]. On much rarer occasions, the high-risk HPV also cause other types of carcinomas [70]. Carcinogenesis is a dead end for HPV because transformed cells produce no infectious virus. Therefore, despite the carcinogenic potential of high-risk HPVs, by and large, these common components of human virome should be considered symbionts, as demonstrated by the protection from skin cancer caused by immunity to ubiquitous skin-infecting HPVs [71] (Figure 1b).

Anelloviruses

Members of the family *Anelloviridae* are among the most enigmatic components of the healthy human virome, both in terms of their evolutionary origin and the impact on human health. Anelloviruses have small icosahedral virions that encapsidate tiny (3 to 3,5 kb) circular ssDNA genomes [72], but unlike the ssDNA viruses in the realm *Monodnaviria*, do not encode recognizable homologs of the rolling circle replication endonuclease. Moreover, no homologs

outside the family have been detected for any of the anellovirus proteins. Hence, anelloviruses are currently not included in the realm *Monodnaviria* and their provenance remains unclear [22,48]. The entire human population is believed to be infected with anelloviruses, and there is no convincing evidence of viral clearance from infected individuals [72,73]. The infections occur at an early age and so far have not been convincingly associated with any disease. The virus load appears to be controlled by the immune system because virus levels increase with the level of host immunosuppression [74,75]. Asymptomatic anellovirus infections are also common in other mammals [73,76-78], suggesting extensive coevolution of anelloviruses with mammalian hosts. Although the potential impact of anelloviruses on human health remains a matter of debate, it has been suggested that they positively influence human physiology by shaping the immunity during early development [72]. Thus, anelloviruses might be the most 'friendly', genuinely symbiotic component of the human virome.

Realm *Duplodnaviria*

The realm *Duplodnaviria* includes viruses with dsDNA genomes that are encapsidated in icosahedral capsids consisting of a distinct type of capsid protein, displaying the HK97 fold (after the first phage for which the capsid protein structure was solved), with the help of the corresponding variety of packaging ATPase, known as the terminase [22]. This realm includes the bulk of the viruses associated with the human microbiome, namely, the numerous tailed bacteriophages, as well as a major component of the virome associated with human cells, the herpesviruses. Some of the human herpesviruses (*Herpesviridae*) infect a variety of cell types and cause life-long latent infections [79]. Of the 9 human herpesviruses identified so far, herpes simplex virus 1 (HSV1), human cytomegalovirus (HCMV), Epstein-Barr virus (EBV) and varicella zoster virus (VZV) are highly prevalent in the human population. Depending on the geographic location, socioeconomic status and, in the case of VZV, vaccination levels, the incidence of these viruses reaches up to 96% [80-82]. The health impact of these viruses (if any) depends on the age and immune system status of the infected person, with ethnicity, gender and genotype being additional significant contributors.

The most stealthy of the human herpes viruses are apparently EBV and HCMV. If acquired in childhood, both of these viruses typically cause asymptomatic infections, mainly, in B-lymphocytes in the case of EBV [83] or in several cell types in the case of HCMV [79]. Such silent infections, however, even if not manifested in disease, cause a range of effects at the molecular,

cellular, tissue and organism levels. In the case of EBV, the gene expression pattern of infected B-cells is reprogrammed [84] and the proportion of plasma cells in blood appears to be increased [7]. Major immunity stimulation effects, a likely result of a balance reached over long virus-host co-evolution process, are well established for HCMV. In particular, ~10% of the $CD4^+$ and $CD8^+$ T cells in latently infected, otherwise healthy adults are HCMV-specific. Likewise, HCMV causes expansion of adaptive $CD57^+$ natural killer (NK) cells targeting virus-infected cells [85]. These powerful arms of the immune system, however, fail to clear the virus due to its armament of immunoevasins, HCMV-encoded proteins involved in modulation of host immunity [86]. Extensive immunity stimulation in HCMV-seropositive individuals appears to enhance responsiveness and protection against heterologous viruses rather than compromise host immunity [86]. These potentially positive effects notwithstanding, damaging consequences of herpesvirus infections appear to be fairly common and vary from mild illness to a variety of life-threatening conditions. Thus, HCMV congenital (*in utero*) infections that occur at up to 2% of childbirths frequently cause neurodevelopmental defects or leukemia [85]. At the other life extremity, in seniors, HCMV seropositivity is associated with increased risks of cancer, cardiovascular disease, and ultimately, mortality.

The delicate, life-long EBV-host balance arising from most childhood infections is also fragile: if infection occurs in even non-immunocompromised young adults, it results in infectious mononucleosis ('kissing disease') [87]. Similarly, EBV is etiologically linked to Hodgkin's and Burkitt's lymphomas and nasopharyngeal carcinoma which are common in Southern Chinese and Eskimo people, as well as to a host of other diseases in immunocompromised individuals [83]. A different pathology pattern is characteristic of VZV infections that cause varicella (chickenpox) in children followed by prolonged latency that, in about 15% cases, leads to virus reactivation in elderly people with weakened immune control, causing zoster (postherpetic neuralgia) [88].

The HSV1 infections exhibit another distinct pattern of virus-host interactions. This virus is normally acquired in childhood causing mainly oral infections; its seroprevalence varies from 70% in developed nations to 100% in developing nations [81,89]. Upon infecting epithelial cells, HSV1 is transmitted to axons and establishes life-long latency in dorsal root ganglia that is periodically manifested in reactivation leading to recurrent acute infections or asymptomatic virus shedding. Similar to other human herpesviruses, HSV1 efficiently evades antiviral innate immune responses mediated by Toll-like and other pathogen recognition receptors. The underlying mechanisms of

HSV1 immunoevasion involve multiple virus proteins targeting diverse innate immunity signaling pathways [90].

Thus, human herpesviruses display a striking variety of cell tropisms and infection patterns that emerged over long-term virus-human co-evolution and co-adaptation. Some herpesviruses, in particular HCMV and EBV, can reach a perfect balance with the human host and often persist for the host's lifetime without causing any pathology. However, because of the complexity of virus-host interactions that involve a variety of genetic, environmental and socioeconomic factors, this balance is fragile and can be broken in many situations resulting in morbidity or even mortality (Figure 1b). Therefore, herpesvirus-host co-existence that involves the majority of the human population blurs the very concept of a 'healthy human virome'.

Realm *Varidnaviria*

This realm includes a broad variety of viruses with dsDNA genomes and icosahedral capsids built of protein unrelated to the capsid proteins of duplodnaviruses [22].Unlike the members of *Duplodnaviria*, varidnaviruses are minor components of the healthy human virome, at best. Several intriguing reports have appeared on the detection of large and giant viruses of the class *Nucleocytoviricota* in healthy humans [91]. Perhaps, the most notable of these findings is the detection in human blood of several members of *Marseilleviridae* one of which has been reported to grow in T lymphocytes [92]. Additionally, several viruses from both *Marseilleviridae* and *Mimiviridae* have been isolated from human stools or detected in human-associated metagenomes [93]. Furthermore, the presence of mimiviruses in peripheral blood mononuclear cells of patients with atypical pneumonia has been reported, and a role for these viruses in the pneumonia pathogenesis has been suggested [94]. Unexpectedly, DNA of Acanthocystis turfacea chlorella virus 1, a member of *Phycodnaviridae*, has been detected in nearly half of the tested oropharyngeal samples from healthy humans, and also has been reported to persist in mouse macrophages [95]. However, so far, many of the reports on the presence of members of *Nucleocytoviricota* in human samples and, especially, their ability to replicate in human cells have been disavowed in follow-up studies [96]. Thus, the status of these viruses as components of the healthy human virome except, perhaps, as occasional contaminants, remains dubious.

Conclusions

The healthy virome of any organism, and especially humans, clearly, is an important component

of the holobiont that makes a major contribution to the health status of the host. However, the very concept of a healthy virome is nebulous and fluid because it is virtually impossible to ascertain that any virus would not cause disease under any conditions. A strong case in point are the herpesviruses that are nearly ubiquitous in the human population, remaining symbionts in most individuals most of the time, but consistently cause disease, in some cases, devastating, in immunocompromised individuals. Conversely, it appears plausible that any virus can become beneficial to the host through protection from other viruses, general stimulation of immunity as in the striking case of β-HPV protecting human hosts from skin cancer, or recruitment of virus genes for host functions. The numerous HERVs integrated in the human genome can be considered the paradigm of virus-host symbiosis. Generally, there is no doubt that many viruses evolved multiple mechanisms to manipulate the host innate and adaptive immunity pathways, ensuring virus persistence and controlling the damage to the host, as most conspicuously exemplified by the latent herpesviruses that are virtually ubiquitous in the human population.

The healthy virome is obviously heterogeneous and consists of 3 distinct components (Figure 1a): i) viruses that systematically enter the human organism, primarily, with food, but do not replicate in humans, ii) viruses infecting prokaryotes and, possibly, unicellular eukaryotes that comprise the healthy human microbiome, and iii) viruses that actually replicate and persist in human cells. With the advances of metagenomics, the human "microbiovirome" has become a subject of intense studies that continue bringing discoveries of new bacteriophage groups. In contrast, the "true" healthy human virome is poorly understood, with many questionable sightings of diverse viruses but little solid evidence on persistence mechanisms. On the whole, and in contrast to the disease-associated virome, the healthy human virome appears to be dominated by DNA viruses, in particular, anelloviruses and herpesviruses, that are substantially more common than RNA viruses in healthy humans. A thorough investigation of this component of the healthy virome can be expected to enhance our understanding of virus-host interactions and have major implications for human health.

Acknowledgements

E.V.K. is supported by the Intramural Research Program of the National Institutes of Health (National Library of Medicine). M.K was supported by l'Agence Nationale de la Recherche grant ANR-20-CE20-0009.

References

1. Koonin EV, Dolja VV: A virocentric perspective on the evolution of life. *Curr Opin Virol* 2013, 3:546-557.
2. Forterre P, Prangishvili D: The major role of viruses in cellular evolution: facts and hypotheses. *Curr Opin Virol* 2013, 3:558-565.
3. Roossinck MJ, Bazan ER: Symbiosis: Viruses as Intimate Partners. *Annu Rev Virol* 2017, 4:123-139.
4. Koonin EV, Makarova KS, Wolf YI, Krupovic M: Evolutionary entanglement of mobile genetic elements and host defence systems: guns for hire. *Nat Rev Genet* 2020, 21:119-131.
5. Rascovan N, Duraisamy R, Desnues C: Metagenomics and the Human Virome in Asymptomatic Individuals. *Annu Rev Microbiol* 2016, 70:125-141.
6. Adiliaghdam F, Jeffrey KL: Illuminating the human virome in health and disease. *Genome Med* 2020, 12:66.

**7. Kumata R, Ito J, Takahashi K, Suzuki T, Sato K: A tissue level atlas of the healthy human virome. *BMC Biol* 2020, 18:55.

The latest broad-scale metagenomics analysis of the healthy human virome.

*8. Shkoporov AN, Clooney AG, Sutton TDS, Ryan FJ, Daly KM, Nolan JA, McDonnell SA, Khokhlova EV, Draper LA, Forde A, et al.: The Human Gut Virome Is Highly Diverse, Stable, and Individual Specific. *Cell Host Microbe* 2019, 26:527-541 e525.

A detailed investigation of the diversity and short scale dynamics of the component of the human healthy virome associated with the gut microbiome.

9. Manrique P, Bolduc B, Walk ST, van der Oost J, de Vos WM, Young MJ: Healthy human gut phageome. *Proc Natl Acad Sci U S A* 2016, 113:10400-10405.
10. Wang H, Ling Y, Shan T, Yang S, Xu H, Deng X, Delwart E, Zhang W: Gut virome of mammals and birds reveals high genetic diversity of the family Microviridae. *Virus Evol* 2019, 5:vez013.
11. Krupovic M, Forterre P: Microviridae goes temperate: microvirus-related proviruses reside in the genomes of Bacteroidetes. *PLoS One* 2011, 6:e19893.
12. Minot S, Bryson A, Chehoud C, Wu GD, Lewis JD, Bushman FD: Rapid evolution of the human gut virome. *Proc Natl Acad Sci U S A* 2013, 110:12450-12455.

*13. Roux S, Krupovic M, Daly RA, Borges AL, Nayfach S, Schulz F, Sharrar A, Matheus Carnevali PB, Cheng JF, Ivanova NN, et al.: Cryptic inoviruses revealed as pervasive in bacteria and archaea across Earth's biomes. *Nat Microbiol* 2019, 4:1895-1906.

An in depth search for inoviruses integrated in bacterial genomes, including those of the human gut microbiota, using the most powerful computational approaches. The results greatly expand the diversity of inovorises and demonstrated their unsuspected high abundance.

14. Dutilh BE, Cassman N, McNair K, Sanchez SE, Silva GG, Boling L, Barr JJ, Speth DR, Seguritan V, Aziz RK, et al.: A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nat Commun* 2014, 5:4498.
15. Yutin N, Makarova KS, Gussow AB, Krupovic M, Segall A, Edwards RA, Koonin EV: Discovery of an expansive bacteriophage family that includes the most abundant viruses from the human gut. *Nat Microbiol* 2018, 3:38-46.

*16. Yutin N, Benler S, Shmakov SA, Wolf YI, Tolstoy I, Rayko M, Antipov D, Pevzner PA, KooninE.V.: Analysis of metagenome-assembled viral genomes from the human gut reveals diverse putative CrAss-like phages with unique genomic features *Nature Communications* 2021, in press.

A comprehensive analysis of virus genomes from the human gut dramatically expanding the diversity of crAss-like phages and demonstrating that this group of phages dominates the human "microbiovirome". Some subgroups of crAss-like phages are shown to possess unique gene organizations, with unprecedented densities of self-splicing introns and intein, and alternative genetic codes.

17. Guerin E, Shkoporov A, Stockdale SR, Clooney AG, Ryan FJ, Sutton TDS, Draper LA, Gonzalez-Tortuero E, Ross RP, Hill C: Biology and Taxonomy of crAss-like Bacteriophages, the Most Abundant Virus in the Human Gut. *Cell Host Microbe* 2018, 24:653-664 e656.

18. Drobysheva AV, Panafidina SA, Kolesnik MV, Klimuk EI, Minakhin L, Yakunina MV, Borukhov S, Nilsson E, Holmfeldt K, Yutin N, et al.: Structure and function of virion RNA polymerase of a crAss-like phage. *Nature* 2020.
19. Norman JM, Handley SA, Baldridge MT, Droit L, Liu CY, Keller BC, Kambal A, Monaco CL, Zhao G, Fleshner P, et al.: Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell* 2015, 160:447-460.
20. Monaco CL, Gootenberg DB, Zhao G, Handley SA, Ghebremichael MS, Lim ES, Lankowski A, Baldridge MT, Wilen CB, Flagg M, et al.: Altered Virome and Bacterial Microbiome in Human Immunodeficiency Virus-Associated Acquired Immunodeficiency Syndrome. *Cell Host Microbe* 2016, 19:311-322.
21. Handley SA, Desai C, Zhao G, Droit L, Monaco CL, Schroeder AC, Nkolola JP, Norman ME, Miller AD, Wang D, et al.: SIV Infection-Mediated Changes in Gastrointestinal Bacterial Microbiome and Virome Are Associated with Immunodeficiency and Prevented by Vaccination. *Cell Host Microbe* 2016, 19:323-335.
**22. Koonin EV, Dolja VV, Krupovic M, Varsani A, Wolf YI, Yutin N, Zerbini FM, Kuhn JH: Global Organization and Proposed Megataxonomy of the Virus World. *Microbiol Mol Biol Rev* 2020, 84.
A comprehensive analysis of the evolutionary relationships among the major groups of viruses. A megataxonomy of viruses is proposed on the basis of these relationships and subsequently adopted by the International Committee for Taxonomy of Viruses (Ref. 23).
23. International Committee on Taxonomy of Viruses Executive C: The new scope of virus taxonomy: partitioning the virosphere into 15 hierarchical ranks. *Nat Microbiol* 2020, 5:668-674.
24. Stapleton JT, Foung S, Muerhoff AS, Bukh J, Simmonds P: The GB viruses: a review and proposed classification of GBV-A, GBV-C (HGV), and GBV-D in genus Pegivirus within the family Flaviviridae. *J Gen Virol* 2011, 92:233-246.
25. Hartlage AS, Cullen JM, Kapoor A: The Strange, Expanding World of Animal Hepaciviruses. *Annu Rev Virol* 2016, 3:53-75.
26. Linnen J, Wages J, Jr., Zhang-Keck ZY, Fry KE, Krawczynski KZ, Alter H, Koonin E, Gallagher M, Alter M, Hadziyannis S, et al.: Molecular cloning and disease association of hepatitis G virus: a transfusion-transmissible agent. *Science* 1996, 271:505-508.
27. Alter MJ, Gallagher M, Morris TT, Moyer LA, Meeks EL, Krawczynski K, Kim JP, Margolis HS: Acute non-A-E hepatitis in the United States and the role of hepatitis G virus infection. Sentinel Counties Viral Hepatitis Study Team. *N Engl J Med* 1997, 336:741-746.
28. Schwarze-Zander C, Blackard JT, Rockstroh JK: Role of GB virus C in modulating HIV disease. *Expert Rev Anti Infect Ther* 2012, 10:563-572.
29. Greenhalgh S, Schmidt R, Day T: Fighting the Public Health Burden of AIDS With the Human Pegivirus. *Am J Epidemiol* 2019, 188:1586-1594.
30. Janowski AB, Krishnamurthy SR, Lim ES, Zhao G, Brenchley JM, Barouch DH, Thakwalakwa C, Manary MJ, Holtz LR, Wang D: Statoviruses, A novel taxon of RNA viruses present in the gastrointestinal tracts of diverse mammals. *Virology* 2017, 504:36-44.
31. Thi Kha Tu N, Thi Thu Hong N, Thi Han Ny N, My Phuc T, Thi Thanh Tam P, Doorn HRV, Dang Trung Nghia H, Thao Huong D, An Han D, Thi Thu Ha L, et al.: The Virome of Acute Respiratory Diseases in Individuals at Risk of Zoonotic Infections. *Viruses* 2020, 12.
32. Wolf YI, Kazlauskas D, Iranzo J, Lucia-Sanz A, Kuhn JH, Krupovic M, Dolja VV, Koonin EV: Origins and Evolution of the Global RNA Virome. *MBio* 2018, 9.
33. Ganesh B, Banyai K, Martella V, Jakab F, Masachessi G, Kobayashi N: Picobirnavirus infections: viral persistence and zoonotic potential. *Rev Med Virol* 2012, 22:245-256.
34. Ganesh B, Masachessi G, Mladenova Z: Animal picobirnavirus. *Virusdisease* 2014, 25:223-238.
*35. Krishnamurthy SR, Wang D: Extensive conservation of prokaryotic ribosomal binding sites in known and novel picobirnaviruses. *Virology* 2018, 516:108-114.
Unexpected observation of highly conserved Shinne-Dalgarno sequences in picobirnaviruses leading to the proposal that these viruses are parasites of prokaryotes.
36. Zhang T, Breitbart M, Lee WH, Run JQ, Wei CL, Soh SW, Hibberd ML, Liu ET, Rohwer F, Ruan Y: RNA viral community in human feces: prevalence of plant pathogenic viruses. *PLoS Biol* 2006, 4:e3.

**37. Symonds EM, Rosario K, Breitbart M: Pepper mild mottle virus: Agricultural menace turned effective tool for microbial water quality monitoring and assessing (waste)water treatment technologies. *PLoS Pathog* 2019, 15:e1007639.
Demonstration of the ubiquity and high abundance of a food-associated plant virus in the human gut.
38. Tokuyama M, Kong Y, Song E, Jayewickreme T, Kang I, Iwasaki A: ERVmap analysis reveals genome-wide transcription of human endogenous retroviruses. *Proc Natl Acad Sci U S A* 2018, 115:12565-12572.
39. Johnson WE: Origins and evolutionary consequences of ancient endogenous retroviruses. *Nat Rev Microbiol* 2019, 17:355-370.
The latest detailed overview of the human endogenous retroviruses. The dual role of retroviruses in animal evolution is emphasized: first, as driveers of defense systems evolution, and second, as reservoir of genetic material for exaptation.
40. Chiu ES, VandeWoude S: Endogenous Retroviruses Drive Resistance and Promotion of Exogenous Retroviral Homologs. *Annu Rev Anim Biosci* 2020.
41. Cullen H, Schorn AJ: Endogenous Retroviruses Walk a Fine Line between Priming and Silencing. *Viruses* 2020, 12.
*42. Alcazer V, Bonaventura P, Depil S: Human Endogenous Retroviruses (HERVs): Shaping the Innate Immune Response in Cancers. *Cancers (Basel)* 2020, 12.
A conceptual article discussing the stimulatory effect of endogenous retroviruses on innate immunity promoting tumor resistance.
43. Koonin EV, Krupovic M: The depths of virus exaptation. *Curr Opin Virol* 2018, 31:1-8.
44. Garcia-Montojo M, Doucet-O'Hare T, Henderson L, Nath A: Human endogenous retrovirus-K (HML-2): a comprehensive review. *Crit Rev Microbiol* 2018, 44:715-738.
45. Xue B, Sechi LA, Kelvin DJ: Human Endogenous Retrovirus K (HML-2) in Health and Disease. *Front Microbiol* 2020, 11:1690.
46. Grandi N, Tramontano E: Human Endogenous Retroviruses Are Ancient Acquired Elements Still Shaping Innate Immune Responses. *Front Immunol* 2018, 9:2039.
*47. Singh M, Cai H, Bunse M, Feschotte C, Izsvak Z: Human Endogenous Retrovirus K Rec forms a Regulatory Loop with MITF that Opposes the Progression of Melanoma to an Invasive Stage. *Viruses* 2020, 12.
This work uncovers the specific mechanism of prevention of the progression of invasive melanoma by the expression of the endogenous viruses of the HERV-K group.
48. Kazlauskas D, Varsani A, Koonin EV, Krupovic M: Multiple origins of prokaryotic and eukaryotic single-stranded DNA viruses from bacterial and archaeal plasmids. *Nat Commun* 2019, 10:3425.
49. Krupovic M, Ghabrial SA, Jiang D, Varsani A: Genomoviridae: a new family of widespread single-stranded DNA viruses. *Arch Virol* 2016, 161:2633-2643.
50. Varsani A, Krupovic M: Smacoviridae: a new family of animal-associated single-stranded DNA viruses. *Arch Virol* 2018, 163:2005-2015.
51. Matthews PC, Sharp C, Simmonds P, Klenerman P: Human parvovirus 4 'PARV4' remains elusive despite a decade of study. *F1000Res* 2017, 6:82.
52. Qiu J, Soderlund-Venermo M, Young NS: Human Parvoviruses. *Clin Microbiol Rev* 2017, 30:43-113.
53. Soderlund-Venermo M: Emerging Human Parvoviruses: The Rocky Road to Fame. *Annu Rev Virol* 2019, 6:71-91.
54. Cotmore SF, Agbandje-McKenna M, Canuti M, Chiorini JA, Eis-Hubinger AM, Hughes J, Mietzsch M, Modha S, Ogliastro M, Penzes JJ, et al.: ICTV Virus Taxonomy Profile: Parvoviridae. *J Gen Virol* 2019, 100:367-368.
55. Allander T: Human bocavirus. *J Clin Virol* 2008, 41:29-33.
56. Guido M, Tumolo MR, Verri T, Romano A, Serio F, De Giorgi M, De Donno A, Bagordo F, Zizza A: Human bocavirus: Current knowledge and future challenges. *World J Gastroenterol* 2016, 22:8684-8697.
57. Ma X, Endo R, Ishiguro N, Ebihara T, Ishiko H, Ariga T, Kikuta H: Detection of human bocavirus in Japanese children with lower respiratory tract infections. *J Clin Microbiol* 2006, 44:1132-1134.
58. Woolf AD, Campion GV, Chishick A, Wise S, Cohen BJ, Klouda PT, Caul O, Dieppe PA: Clinical manifestations of human parvovirus B19 in adults. *Arch Intern Med* 1989, 149:1153-1156.

59. Bonvicini F, La Placa M, Manaresi E, Gallinella G, Gentilomi GA, Zerbini M, Musiani M: Parvovirus b19 DNA is commonly harboured in human skin. *Dermatology* 2010, 220:138-142.
60. Moens U, Krumbholz A, Ehlers B, Zell R, Johne R, Calvignac-Spencer S, Lauber C: Biology, evolution, and medical importance of polyomaviruses: An update. *Infect Genet Evol* 2017, 54:18-38.
61. Spriggs CC, Laimins LA: Human Papillomavirus and the DNA Damage Response: Exploiting Host Repair Pathways for Viral Replication. *Viruses* 2017, 9.
62. Buck CB, Van Doorslaer K, Peretti A, Geoghegan EM, Tisza MJ, An P, Katz JP, Pipas JM, McBride AA, Camus AC, et al.: The Ancient Evolutionary History of Polyomaviruses. *PLoS Pathog* 2016, 12:e1005574.
63. Krump NA, Liu W, You J: Mechanisms of persistence by small DNA tumor viruses. *Curr Opin Virol* 2018, 32:71-79.
64. Cortese I, Reich DS, Nath A: Progressive multifocal leukoencephalopathy and the spectrum of JC virus-related disease. *Nat Rev Neurol* 2020.
65. Krump NA, You J: Molecular mechanisms of viral oncogenesis in humans. *Nat Rev Microbiol* 2018, 16:684-698.
66. Doorbar J, Egawa N, Griffin H, Kranjec C, Murakami I: Human papillomavirus molecular biology and disease association. *Rev Med Virol* 2015, 25 Suppl 1:2-23.
67. Graham SV: The human papillomavirus replication cycle, and its links to cancer progression: a comprehensive review. *Clin Sci (Lond)* 2017, 131:2201-2221.
68. Chiu YF, Sugden B: Plasmid Partitioning by Human Tumor Viruses. *J Virol* 2018, 92.
69. Serrano B, Brotons M, Bosch FX, Bruni L: Epidemiology and burden of HPV-related disease. *Best Pract Res Clin Obstet Gynaecol* 2018, 47:14-26.
70. Araldi RP, Sant'Ana TA, Modolo DG, de Melo TC, Spadacci-Morena DD, de Cassia Stocco R, Cerutti JM, de Souza EB: The human papillomavirus (HPV)-related cancer biology: An overview. *Biomed Pharmacother* 2018, 106:1537-1556.
**71. Strickley JD, Messerschmidt JL, Awad ME, Li T, Hasegawa T, Ha DT, Nabeta HW, Bevins PA, Ngo KH, Asgari MM, et al.: Immunity to commensal papillomaviruses protects against skin cancer. *Nature* 2019, 575:519-522.
Squamous cell skin carcinoma has been associated with the replication of human b-HPV. This work shows that such correlation does not reflect causative role of b-HPV in cancer progression, but on the contrary, occurs because T call mediated immunity against b-HPV suppresses tumorigenesis.
**72. Kaczorowska J, van der Hoek L: Human anelloviruses: diverse, omnipresent and commensal members of the virome. *FEMS Microbiol Rev* 2020, 44:305-313.
The latest review on anelloviruses emphasizing the ubiquity of these viruses in humans, the lack of convincing disease association and the potential beneficial roles.
73. Webb B, Rakibuzzaman A, Ramamoorthy S: Torque teno viruses in health and disease. *Virus Res* 2020, 285:198013.
74. Focosi D, Antonelli G, Pistello M, Maggi F: Torquetenovirus: the human virome from bench to bedside. *Clin Microbiol Infect* 2016, 22:589-593.
75. Blatter JA, Sweet SC, Conrad C, Danziger-Isakov LA, Faro A, Goldfarb SB, Hayes D, Jr., Melicoff E, Schecter M, Storch G, et al.: Anellovirus loads are associated with outcomes in pediatric lung transplantation. *Pediatr Transplant* 2018, 22.
76. Fahsbender E, Burns JM, Kim S, Kraberger S, Frankfurter G, Eilers AA, Shero MR, Beltran R, Kirkham A, McCorkell R, et al.: Diverse and highly recombinant anelloviruses associated with Weddell seals in Antarctica. *Virus Evol* 2017, 3:vex017.
77. de Souza WM, Fumagalli MJ, de Araujo J, Sabino-Santos G, Jr., Maia FGM, Romeiro MF, Modha S, Nardi MS, Queiroz LH, Durigon EL, et al.: Discovery of novel anelloviruses in small mammals expands the host range and diversity of the Anelloviridae. *Virology* 2018, 514:9-17.
78. Crane A, Goebel ME, Kraberger S, Stone AC, Varsani A: Novel anelloviruses identified in buccal swabs of Antarctic fur seals. *Virus Genes* 2018, 54:719-723.
79. Connolly SA, Jardetzky TS, Longnecker R: The structural basis of herpesvirus entry. *Nat Rev Microbiol* 2020.
80. Cohen JI: Epstein-Barr virus infection. *N Engl J Med* 2000, 343:481-492.
81. Gupta R, Warren T, Wald A: Genital herpes. *Lancet* 2007, 370:2127-2137.

82. Zuhair M, Smit GSA, Wallis G, Jabbar F, Smith C, Devleesschauwer B, Griffiths P: Estimation of the worldwide seroprevalence of cytomegalovirus: A systematic review and meta-analysis. *Rev Med Virol* 2019, 29:e2034.
83. Taylor GS, Long HM, Brooks JM, Rickinson AB, Hislop AD: The immunology of Epstein-Barr virus-induced disease. *Annu Rev Immunol* 2015, 33:787-821.
84. Moustafa A, Xie C, Kirkness E, Biggs W, Wong E, Turpaz Y, Bloom K, Delwart E, Nelson KE, Venter JC, et al.: The blood DNA virome in 8,000 humans. *PLoS Pathog* 2017, 13:e1006292.
85. Semmes EC, Hurst JH, Walsh KM, Permar SR: Cytomegalovirus as an immunomodulator across the lifespan. *Curr Opin Virol* 2020, 44:112-120.
**86. Berry R, Watson GM, Jonjic S, Degli-Esposti MA, Rossjohn J: Modulation of innate and adaptive immunity by cytomegaloviruses. *Nat Rev Immunol* 2020, 20:113-127.
A comprehensive discussion of immunoevasins, the diverse proteins encoded by human cytomegaloviruses involved in the evasion and modulation of the host immune system.
87. Cruz-Munoz ME, Fuentes-Panana EM: Beta and Gamma Human Herpesviruses: Agonistic and Antagonistic Interactions with the Host Immune System. *Front Microbiol* 2017, 8:2521.
88. Gershon AA, Breuer J, Cohen JI, Cohrs RJ, Gershon MD, Gilden D, Grose C, Hambleton S, Kennedy PG, Oxman MN, et al.: Varicella zoster virus infection. *Nat Rev Dis Primers* 2015, 1:15016.
89. Saleh D, Yarrarapu SNS, Sharma S: Herpes Simplex Type 1. In *StatPearls*. Edited by; 2020.
90. Zhu H, Zheng C: The Race between Host Antiviral Innate Immunity and the Immune Evasion Strategies of Herpes Simplex Virus 1. *Microbiol Mol Biol Rev* 2020, 84.
91. Colson P, Aherfi S, La Scola B, Raoult D: The role of giant viruses of amoebas in humans. *Curr Opin Microbiol* 2016, 31:199-208.
92. Popgeorgiev N, Boyer M, Fancello L, Monteil S, Robert C, Rivet R, Nappez C, Azza S, Chiaroni J, Raoult D, et al.: Marseillevirus-like virus recovered from blood donated by asymptomatic humans. *J Infect Dis* 2013, 208:1042-1050.
93. Colson P, Fancello L, Gimenez G, Armougom F, Desnues C, Fournous G, Yoosuf N, Million M, La Scola B, Raoult D: Evidence of the megavirome in humans. *J Clin Virol* 2013, 57:191-200.
94. Colson P, La Scola B, Raoult D: Giant viruses of amoebae as potential human pathogens. *Intervirology* 2013, 56:376-385.
95. Petro TM, Agarkova IV, Zhou Y, Yolken RH, Van Etten JL, Dunigan DD: Response of Mammalian Macrophages to Challenge with the Chlorovirus Acanthocystis turfacea Chlorella Virus 1. *J Virol* 2015, 89:12096-12107.
96. Abrahao J, Silva L, Oliveira D, Almeida G: Lack of evidence of mimivirus replication in human PBMCs. *Microbes Infect* 2018, 20:281-283.

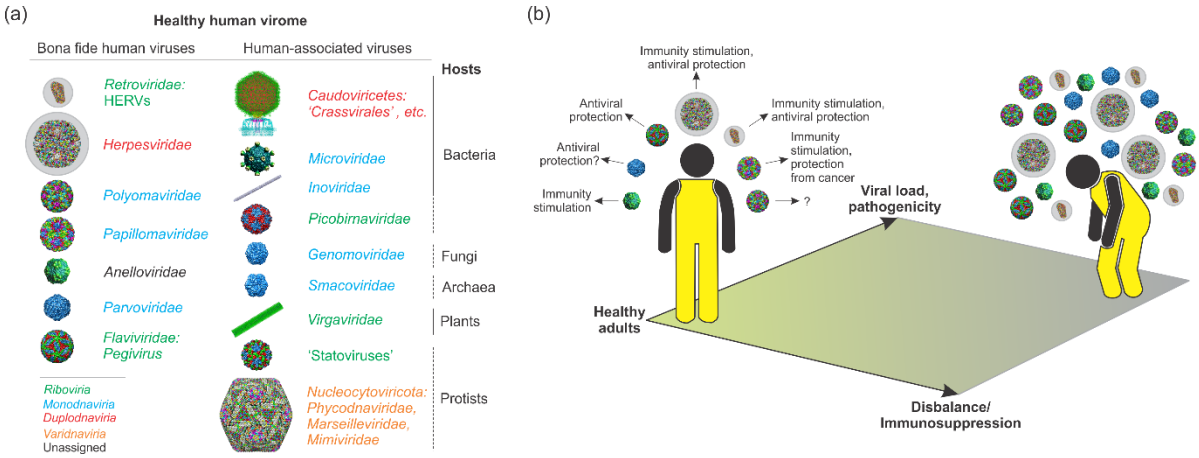

Figure 1. The healthy human virome.

(a) Taxa of viruses found in healthy humans. Taxa of viruses replicating in human cells are shown on the left, whereas those of viruses infecting human-associated microbes or associated with food sources are shown on the right (the actual or suspected hosts are listed next to the corresponding taxa, with the uncertain assignments indicated with broken lines). Depicted virus taxa are represented with virion structures which were retrieved from VIPERdb (viperdb.scripps.edu) or the Electron Microscopy Data Bank (https://www.ebi.ac.uk/pdbe/emdb/). When the structure of a virus representing a particular taxon was not available, a structurally related member was chosen instead. Virus taxa are colored according to their realm affiliation (the key is provided at the bottom).

(b) Fluidity of the healthy human virome. The (potential) beneficial effects of the healthy human virome are indicated next to the corresponding virus structures. Question marks denote uncertainty. Upon changes in the health status/immunosuppression, viruses that cause asymptomatic infections or are beneficial in healthy individuals proliferate and can cause diseases including severe ones. The figure illustrates the general tendency of increased virus load under immunosuppression/disease conditions and should not be interpreted as a quantitative representation of the changes for any depicted group of viruses.