



ACADÉMIE
DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus

Biologies


Mart Krupovic and Eugene V. Koonin

Organization and evolution of the virosphere and the replicator space

Volume 349 (2026), p. 35-75

Online since: 11 May 2026

<https://doi.org/10.5802/crbiol.193>

 This article is licensed under the
CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.
<http://creativecommons.org/licenses/by/4.0/>



*The Comptes Rendus. Biologies are a member of the
Mersenne Center for open scientific publishing*
www.centre-mersenne.org — e-ISSN : 1768-3238



Review article

Organization and evolution of the virosphere and the replicator space

Mart Krupovic^{✉,*,#,a} and Eugene V. Koonin^{✉,*,#,b}

^a Institut Pasteur, Université Paris Cité, Cell Biology and Virology of Archaea Unit, Paris, France

^b Computational Biology Branch, Division of Intramural Research, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA

E-mails: mart.krupovic@pasteur.fr (M. Krupovic), koonin@ncbi.nlm.nih.gov (E. V. Koonin)

Abstract. Viruses are obligate symbionts of cellular life forms that can replicate only within host cells and typically form virions (virus particles) to spread among host organisms. Virions numerically dominate the biosphere, exceeding the number of cells several-fold, and also comprise the main reservoir of genetic diversity on earth. Nearly all organisms host multiple, diverse viruses. Unlike cellular organisms, viruses have genomes (genetic information carriers incorporated into virions) that consist of all forms of RNA and DNA, suggesting an evolutionary connection between extant viruses and the primordial replicator pool. Lately, extensive mining of metagenomes and metatranscriptomes has dramatically expanded the world of viruses (virosphere), revealing an unsuspected and unprecedented diversity. Viruses share no universal genes and have multiple origins. However, about 15 viral hallmark genes each bring together multiple, diverse groups of viruses, and many other genes are shared within such groups. Evolution of viruses is inextricably intertwined with the evolution of their hosts. A key aspect of virus–host coevolution is the arms race resulting in accelerated evolution on both sides, especially of host defenses and viral counter-defenses. A complementary, prominent feature of this coevolution is exaptation, whereby viral genes are coopted by the hosts for antiviral defense and other roles, and conversely, viruses capture host genes for diverse functions in virus replication, virion morphogenesis and virus–host interaction. In this review, we attempt a synthesis of the current understanding of the global organization of the virosphere, the major trends and events in the evolution of viruses, and the high-level taxonomy of viruses.

Keywords. Virosphere, Evolution and origins of viruses, Viral hallmark genes, Virus–host coevolution, Virus megataxonomy.

Note. Article submitted by invitation.

Manuscript received 3 February 2026, revised and accepted 1 April 2026, online since 11 May 2026.

This article is dedicated to the memory of the outstanding virologist Dennis H. Bamford (1948–2025)

1. Introduction

Viruses are best known as agents of human, animal and plant diseases, some of these widespread

and devastating. Suffice it to mention numerous epidemics of smallpox prior to the introduction of broad vaccination, the 1918–1920 pandemic of Spanish flu, the catastrophic epidemics of poliomyelitis in the 1930s and 1940s, eventually curtailed by vaccines, and obviously, the recent COVID-19 pandemic (Rochman et al., 2022). Other viruses, such as certain

* Corresponding authors

Contributed equally

strains of human papillomaviruses or Epstein–Barr virus, are known to cause cancer (zur Hausen, 2002; Shannon-Lowe and Rickinson, 2019). Relatively well known are also viruses causing disease in plants, such as the tobacco mosaic virus, the very first virus that was discovered in the late 19th century (Crea-ger, 2022). However, it is much less commonly appreciated that viruses causing diseases in animals and plants, however numerous and diverse, are but tiny parts of the vast virus world, also known as the virosphere. The myriads of other viruses comprising the virosphere infect virtually all other organisms on the planet, in many cases killing the host, but in many others, causing no harm or even being beneficial.

Viruses are defined as mobile genetic elements (MGE) encoding at least one protein that is a major component of virions (virus particles) encasing the genome of the virus (Koonin, Dolja, Krupovic and Kuhn, 2021). Viruses are obligate intracellular symbionts that reproduce inside cells of all organisms, with the sole possible exception of some of the most drastically reduced intracellular bacterial symbionts of eukaryotes. Moreover, most organisms, from bacteria to humans, are hosts to a broad variety of viruses with a wide range of genome sizes, diverse cycles of genome replication and expression, virion structures and modes of interaction with the hosts. When describing viruses, we purposefully use the umbrella term “symbionts” rather than the more commonly used “parasites” because virus–host interactions span the entire spectrum of symbiotic relationships with the hosts, from aggressive parasitism to commensalism to mutualism (Gonzalez et al., 2020).

The emergence of MGE, including viruses, is an intrinsic, fundamentally inevitable, and essential feature of the evolution of life (Koonin, Wolf, et al., 2017). There are strong theoretical arguments and overwhelming empirical evidence that MGE, including, most likely, bona fide viruses, have been associated with the evolution of cellular life on Earth ever since its earliest stages, that is, for more than 4 billion years (Koonin, Senkevich, et al., 2006; Koonin and Dolja, 2013; Forterre and Prangishvili, 2009; Forterre and Prangishvili, 2013). In themselves, viruses cannot be legitimately considered life forms inasmuch as they are not reproducers, that is, they lack physical continuity across generations, and strictly depend on the host cells for their replication (Moreira

and Lopez-Garcia, 2009; Lopez-Garcia, 2012). Instead, viruses are replicators, maintaining genetic continuity and autonomy, and following evolutionary trajectories that are distinct from, even if entangled with, those of the hosts (Koonin, Dolja, Krupovic and Kuhn, 2021).

In the last decade, our understanding of the world of viruses has been transformed through the combination of the explosive growth of metagenomic and metatranscriptomic sequence databases, and sustained efforts on the discovery of novel viruses and the elucidation of the evolutionary relationships within the virosphere (Simmonds, Adams, et al., 2017; Koonin, Dolja, Krupovic, Varsani, et al., 2020; Roux and Coclet, 2026). These developments that take advantage of advanced methods for sequence and protein structure analysis have been codified in the new, comprehensive evolutionary taxonomy of viruses that has been formally adopted by the International Committee on Taxonomy of Viruses (ICTV) (Simmonds, Adriaenssens, Lefkowitz, et al., 2025). In this review article, we undertake to present a synthesis of these revolutionary advances, outlining the global structure of the virosphere along with the concepts, trends and specific scenarios of the evolution of the major groups of viruses. We further explore the connections between viruses and other types of MGE, such as plasmids and transposons, and the position of the virosphere within the replicator space.

2. The dimensions of the virosphere

Before discussing our current understanding of the evolution of the virosphere, we outline its key characteristics and dimensions. Unlike cellular life forms, which all share the same fundamental scheme of genetic information storage and transmission, with genomes consisting of enormous double-stranded (ds) DNA molecules, viruses explore the entire space of possible nucleic acid-based information storage and transmission route cycles (Baltimore, 1971; Agol, 1974; Koonin, Krupovic and Agol, 2021). All forms of DNA and RNA are used as genomes, that is, the information carriers incorporated into virions, by different groups of viruses. The nature of the genome defines the viral replication-expression cycle, and accordingly viruses are traditionally divided into 7 so-called Baltimore classes (first introduced in David Baltimore’s seminal 1971 paper

(Baltimore, 1971)) (Figure 1A). Within different Baltimore classes, viral genomes can be either circular or linear and either continuous or segmented. The size of viral genomes also depends on the Baltimore class. Viruses of 6 of the 7 classes, with the exception of class I (dsDNA genomes), have (relatively) small genomes, within the range of 1 to about 60 kilobases (kb). By contrast, viral dsDNA genomes span 3 orders of magnitude from about 5 kb to more than 4 megabases (Mb) (Figure 1B,C). Notably, the largest genomes of the so-called giant viruses exceed in size and the number of genes the genomes of numerous bacteria and archaea (and some parasitic eukaryotes as well), defying the notion of viruses as minuscule entities and obliterating the boundary between viruses and cellular life forms with regard to genetic complexity (Claverie, 2006; Koonin and Yutin, 2019).

The diversity of the molecular structures and the wide size range of viral genomes are complemented by the diversity of sizes, shapes and molecular organizations of virions. Most virions are small, but like in the case of the genomes, the particles of giant viruses are larger than many bacterial and archaeal cells. In the great majority of virions, the core structure is the capsid, the proteinaceous shell encasing the genome. Most common capsids adopt one of the few shapes, primarily, spherical or filamentous, built on icosahedral or helical symmetry, respectively (Crick and Watson, 1956; Caston and Carrascosa, 2013). There is, however, a broad variety of comparatively rare, odd capsid shapes as well as several groups of viruses that lack a typical capsid while retaining a glycoprotein-rich lipid envelope encasing the genomic nucleic acid (see below). Furthermore, some MGE that are traditionally considered viruses lack genes for structural proteins and do not package their genomes at all, while having clearly evolved from typical, capsid-encoding viruses (Koonin and Dolja, 2014) (see also discussion below). In some viruses, virions are multilayered, consisting of more than one icosahedral shell, whereas in other viruses, the protein capsid is enveloped by a lipid membrane and/or contains an internal membrane surrounding the genome.

A central theme in the study of virus evolution is the origin of functional and structural features by divergence from common ancestors vs. independent, convergent emergence. In the discussion

that follows, we shall see that neither the Baltimore classes of viruses nor the basic shapes and organizations of virions are monophyletic, convergence being a major trend in the evolution of the virosphere. To a large extent, evolutionary convergence is underpinned by physical constraints. As a major case in point, the number of basic capsid shapes is limited to only a few, with the great majority of capsids being either spherical or filamentous (often described as rod-shaped in the case of rigid filaments), which are two simple, symmetrical, thermodynamically favorable architectures. A larger variety of odd capsid shapes are rare in the virosphere. Importantly, convergence at the level of capsid shapes or morphologies does not translate to convergence of the actual molecular solutions to capsid building. Indeed, both icosahedral and helical capsids are constructed from capsid proteins with multiple distinct structural folds (Krupovic and Koonin, 2017b). Conversely, there is no evidence that structurally similar capsid proteins convergently evolved from structurally distinct ancestral proteins. Thus, structural similarity among capsid proteins within a particular virus group is interpreted as evidence of common ancestry and is one of the cornerstones of megataxonomy. In a more general context, although viruses are clearly polyphyletic, origin of new types of viruses is rare and is typically followed by extensive evolutionary diversification, as discussed below.

There is a popular meme that virus particles are the most abundant biological entities on Earth. The number of virions present on the planet at any given time has been estimated at the hyper-astronomical value of about 10^{31} , with a virus/microbe ratio (VMR) of 10 or higher (Hendrix, 2003; Mushegian, 2020). However, these striking VMR values have been calculated largely by using epifluorescence microscopy or flow cytometry to count extracellular virions, primarily, in marine environments. Both these techniques are prone, on the one hand, to missing intracellular virions (including lysogens), but on the other hand, to erroneously reporting non-viral particles as virions (Forterre, Soler, et al., 2013). More reliable estimates for DNA viruses have been obtained recently by quantification of the major capsid protein genes in metagenomes compared to host hallmark genes (namely, ribosomal protein genes). These analyses yielded a more nuanced picture whereby VMR was generally lower than previously thought,

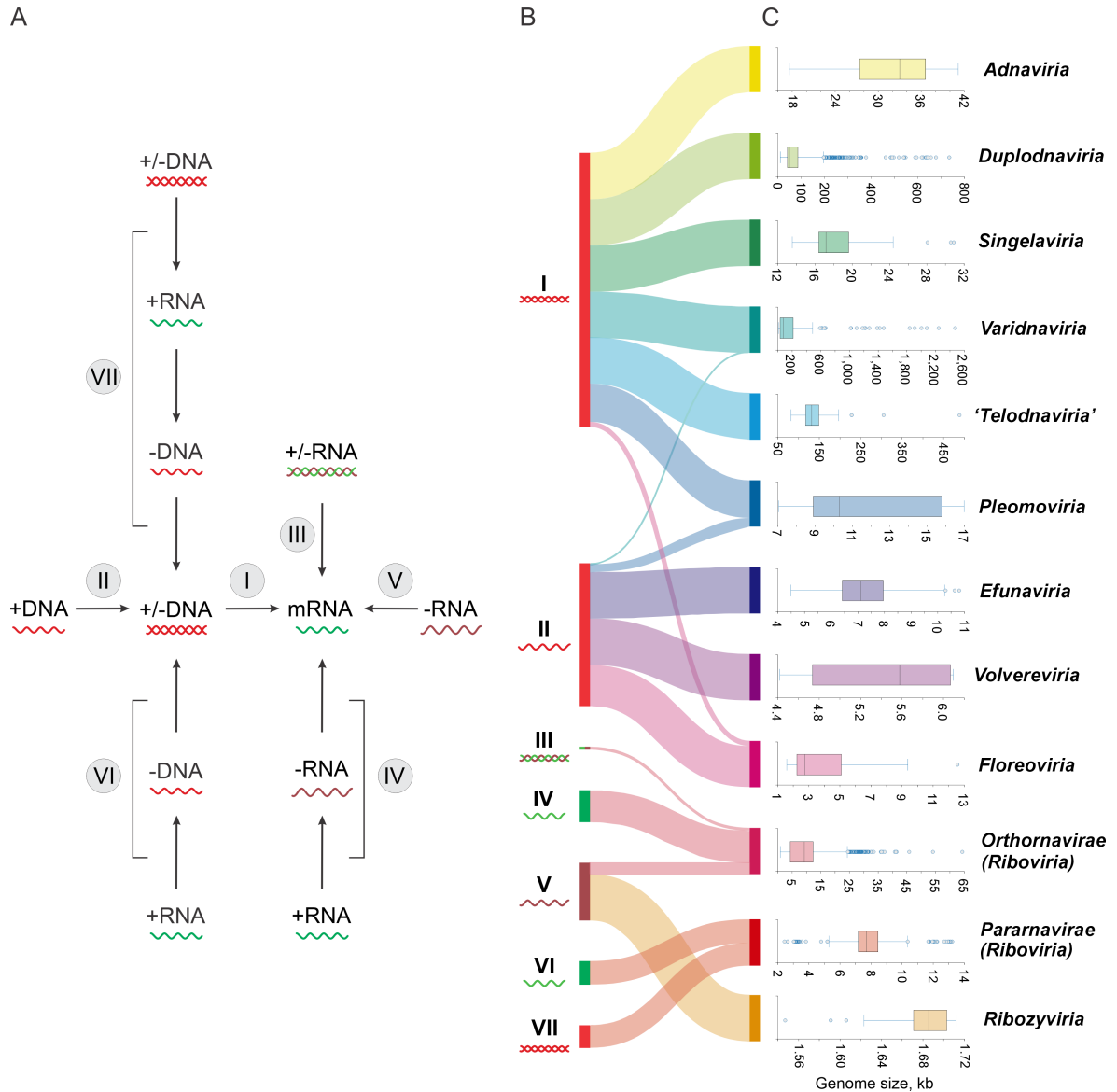


Figure 1. The Baltimore classes and realms of viruses. (A) Replication and expression of viral genomes in the 7 Baltimore classes. Each of the classes is defined according to the form of the nucleic acid incorporated into virions (the viral genome). (B) Sankey diagram showing the correspondence between the Baltimore classes and viral realms. (C) Range of genome sizes (kb) in each viral realm. The genome size data for classified representatives of each realm was retrieved from the NCBI GenBank database (Sayers et al., 2025).

about 2 on average, and broadly varied across environments (Lopez-Garcia, Gutierrez-Preciado, et al., 2023). The highest VMR, 3–4 on average, was observed in metagenomes from aquatic environments, whereas in animal-associated metagenomes, the typical VMR was much lower, with the mean value around 1 (ibid.), in agreement with previous esti-

mates for bacterial (Carding et al., 2017; Shkoporov and Hill, 2019) and archaeal (Baquero, Medvedeva, et al., 2024) viruses. Intermediate VMR values were reported for soils, sediments, and microbial mats. These findings would bring down the estimates of virus abundance in the biosphere by about an order of magnitude, but still support the excess of virus

particles over cells on both the planetary scale and in most habitats. The VMR is not a linear function of the microbial density in the given environment. Analysis of thousands of estimates from marine samples demonstrated that the dependence of the VMR on microbial density is best described by power law functions with exponents below unity, that is, the ratio of virions to cells decreases with increased microbial population density (Wigington et al., 2016). At least in part, this dependency is explained by the piggyback-the-winner model for temperate bacteriophages whereby lysogeny becomes advantageous for a virus compared to lytic growth at high host density (B. Knowles et al., 2017).

Although a comprehensive census of the virosphere remains a major challenge for the future, simple, ballpark estimates can give us an idea of the total number of virus species and viral genes in the biosphere (Koonin, Krupovic and Dolja, 2023). The bulk of the diversity in the virosphere is accounted for by tailed bacteriophages with large dsDNA genomes, with the other phages, archaeal and eukaryotic viruses adding relatively little. There are at least 10^6 – 10^7 species of bacteria in the biosphere, with some estimates suggesting orders of magnitude more. A typical bacterial species is the host to multiple viruses. For example, for the *Escherichia coli* strain K12, over 100 distinct viruses are currently known (Humolli et al., 2025), whereas for a single strain of *Mycobacterium smegmatis*, more than 2000 viruses have been identified following extensive sampling (Hatfull, 2020; Abad et al., 2023; Hatfull, 2022). In all likelihood, these numbers underestimate the actual diversity of the viromes of the respective bacterial species. Conservatively assuming 10–100 viral species per bacterial host species, the global virome can be estimated to include at least 10^7 – 10^9 virus species. The upper bound on this range should be considered a more realistic estimate, given the conservative assumptions used. Currently, only about 17 000 virus species are formally recognized (Simmonds, Adriaenssens, Lefkowitz, et al., 2025), and even taking into account numerous new ones that are waiting to be established based on the results of metagenome and metatranscriptome mining, only a tiny part of the virosphere has been explicitly described so far. However, as we shall see below, this small sampling is likely to be representative of the large, distinct regions of the virosphere.

Based on the simple calculations above, the genetic diversity of the virosphere, that is, the total number of unique genes in the global virome, can be roughly estimated as well (Koonin, Krupovic and Dolja, 2023). The large dsDNA genomes of typical bacteriophages that quantitatively dominate the virosphere encompass many lineage-specific genes that are mostly involved in counteracting the host antiviral defenses (Yan et al., 2023). Conservatively assuming 10 such unique genes per virus species, the plausible low bound of the virosphere diversity can be estimated at about 10^{10} genes, which is likely one to two orders of magnitude greater than the number of unique genes in cellular life forms. Thus, the virosphere is the main reservoir of genetic novelty on our planet. These estimates are generally consistent with the more formal attempt to estimate the potential virus genome and protein space by simulating the process of virus discovery in viral metagenomic studies. Using the power function, which was found to best fit the increasing trends of virus diversity, it was estimated that there are at least 8.2×10^8 viral operational taxonomic units and 1.6×10^9 viral protein clusters on Earth (Lu et al., 2025). These estimates suggest that less than 3% of the viral genetic diversity has been uncovered thus far.

The general message from this brief survey of the virosphere dimensions is that, although viruses are in general tiny compared to cellular life forms and fully depend on the hosts for their replication, the virosphere is vast and enormously diverse, in some dimensions, exceeding the diversity of cellular organisms. Understanding the evolution of the biosphere without deep exploration of the virosphere would be a hopeless undertaking.

3. The metagenomic revolution and virosphere expansion

In the last decade, the knowledge of virus diversity, which can be translated into reconstruction of the virosphere evolution, took a major leap forward thanks to the advances of metagenomics and metatranscriptomics (Simmonds, Adams, et al., 2017). In the past, the study of viruses required propagation in the laboratory, either in the actual host organisms or in cell culture. Accordingly, the exploration of the virosphere diversity had been limited to the relatively few species of well-studied animals and

plants and several model protist, bacterial and archaeal species. However, it is well known from the surveys of microbial diversity by 16S RNA sequencing that only about 0.1% of microbes currently can be grown in the laboratory (Rinke et al., 2013; Solden et al., 2016). Thus, the traditional approaches only scratch the proverbial surface of the global microbiome and, consequently, of the global virome. The fast progress of metagenomic and metatranscriptomic approaches has radically changed this situation (Koonin, Makarova and Wolf, 2021). Indeed, already in 2017, metagenome and metatranscriptome mining became by far the most important source of new virus discovery (Simmonds, Adams, et al., 2017). It is difficult to come up with exact estimates, but clearly, the number of viruses representing putative distinct virus species discovered by metagenome and metatranscriptome mining now exceeds that identified by traditional laboratory methods by several orders of magnitude. The increasingly detailed understanding of the global organization and evolution of the virosphere that we discuss below could not have been possibly attained without extensive analysis of metagenomic and metatranscriptomic sequence databases. Importantly, acknowledging the dominant role of metagenomics and metatranscriptomics in the exploration of the virosphere, the ICTV in 2017 formally recognized the legitimacy of viral species and higher taxa established based solely on sequence analysis, without isolation of the actual virus (Simmonds, Adams, et al., 2017; Adams et al., 2017).

4. The global structure and evolution of the virosphere and viral megataxonomy

4.1. Viral hallmark genes

A fundamental fact that needs to be emphasized before even starting to discuss the structure and evolution of the virosphere is that viruses do not share a single common origin, in sharp contrast to cellular life forms. About 100 genes are strictly universal among cellular organisms and can be used to build the “Tree of life”, individually or more commonly, in different combinations (Koonin, 2003; Ciccarelli et al., 2006). Viruses, however, lack any universal genes, and accordingly, there can be no single tree of viruses, in principle (Koonin, Senkevich, et al., 2006). Put another way, unlike cellular life

forms, which all can be traced back to the Last Universal Cellular Ancestor (LUCA), viruses are polyphyletic, that is, had multiple origins. That said, although the total number of independent viral origins is hard to estimate, there are relatively few vast assemblages of viruses for which common ancestry is traceable.

The evolutionary connections between diverse large groups of viruses were traced by comparative and phylogenetic analysis of shared viral hallmark genes (VHG) (Iranzo, Krupovic, et al., 2016; Iranzo, Koonin, et al., 2016; Wolf, Kazlauskas, et al., 2018; Koonin, Dolja and Krupovic, 2022). The VHGs are difficult to define formally because of an apparent circularity inherent in that definition: VHGs are conserved across diverse groups of viruses which are themselves defined through shared VHGs. Informally, however, VHGs are readily identifiable, and their identities are not surprising: these are genes encoding major virion morphogenesis (primarily, capsid) proteins and key components of viral replication machineries (Table 1). In viruses with small genomes, the VHGs represent a large fraction of the genes, and in some cases, all of the genes. By contrast, in viruses with large genomes, the VHGs can comprise less than 1% of the genes. Nevertheless, in all viruses, the VHGs are central both to virus replication and morphogenesis, and to the study of the evolution of the virosphere.

The VHGs define the distinct domains of the virosphere that correspond to viral realms, the top rank in virus taxonomy (International Committee on Taxonomy of Viruses Executive Committee, 2020). In accord with the fundamental principles of evolutionary taxonomy, each realm is supposed to be monophyletic, that is, all viruses in a realm are assumed to share a common origin (Koonin, Dolja, Krupovic, Varsani, et al., 2020; Simmonds, Adriaenssens, Zerbini, et al., 2023). In actuality, this common origin can be limited to sharing only a few or even a single VHG. Nevertheless, given the central roles of the VHGs in virus replication, this approach is productive for reconstructing the evolution of the virosphere and delineating higher taxa of viruses to develop the taxonomic system that is informally known as megataxonomy (Koonin, Dolja, Krupovic, Varsani, et al., 2020; Koonin, Kuhn, et al., 2024). By the conventions of taxonomy, any virus can belong to only one realm although, when some viruses combine

Table 1. Virus hallmark proteins

Protein	Structural fold	Function	Distribution in the virosphere
RNA-dependent RNA polymerase	RRM	Replication and expression of RNA genomes	<i>Orthornavirae</i>
Reverse transcriptase	RRM	Replication of all RNA and DNA genomes with a reverse transcription stage in the replication cycle	<i>Pararnavirae</i>
HUH superfamily rolling circle replication endonuclease	RRM	Initiation of rolling circle DNA replication	<i>Efunaviria</i> , <i>Volvereviria</i> , <i>Floreoviria</i> , <i>Pleomoviria</i> , some <i>Duplodnaviria</i> , <i>Varidnaviria</i> and <i>Adnaviria</i>
Family B DNA polymerase	RRM	Replication of dsDNA genomes	Many <i>Duplodnaviria</i> , <i>Varidnaviria</i> , <i>Naldaviricetes</i> , <i>Bidnaviridae</i> , several families of archaeal viruses
Archaeo-eukaryotic primase	RRM	Priming of DNA synthesis	Many <i>Duplodnaviria</i> , <i>Varidnaviria</i> , <i>Naldaviricetes</i>
Superfamily 3 helicase	P-loop NTPase (AAA+ ATPase class)	Unwinding of RNA and DNA genomes during replication	Many <i>Riboviria</i> , <i>Duplodnaviria</i> , <i>Varidnaviria</i> , <i>Floreoviria</i> , <i>Naldaviricetes</i>
FtsK-like ATPase	P-loop NTPase	ATP-dependent genome packaging	Most <i>Varidnaviria</i> , <i>Singelaviria</i> , <i>Efunaviria</i>
Terminase, large subunit	P-loop NTPase (RecA class)—RNase H-fold nuclease	ATP-dependent genome packaging and processing	<i>Duplodnaviria</i>
Single jelly-roll capsid protein	Jelly-roll (8-stranded β -barrel)	Capsid formation: major and minor capsid proteins	<i>Volvereviria</i> , <i>Floreoviria</i> , <i>Orthornavirae</i> , <i>Singelaviria</i> , <i>Varidnaviria</i>
Double jelly-roll capsid protein	Double jelly-roll (2 tandem 8-stranded β -barrels)	Capsid formation: major capsid protein	Majority of <i>Varidnaviria</i>
HK97-like capsid protein	HK97-fold	Capsid formation: major capsid protein	<i>Duplodnaviria</i>
Portal protein	Unique, mostly α -helical, “portal” fold	Capsid assembly and genome packaging	<i>Duplodnaviria</i>

VHGs characteristic of different realms, an apparent conflict emerges. Below we discuss some such cases.

Initially, four major realms were established, *Riboviria*, *Monodnaviria*, *Varidnaviria*, and *Duplodnaviria*, each including a vast diversity of viruses, and two much smaller realms were added shortly thereafter. However, subsequent in-depth evolutionary analysis led to the split of *Varidnaviria* into two realms, and a split of *Monodnaviria* into four realms, with the new realms approved by ICTV. More splitting of realms based on a stricter approach to monophyly can be expected, as discussed below.

4.2. *Riboviria: the expanding viral RNA World*

The viruses in the realm *Riboviria* share a single VHG, namely, the gene encoding RNA-directed RNA polymerases (RdRPs) or RNA-directed DNA polymerases (reverse transcriptases, RTs), the homologous enzymes responsible for viral genome replication and expression (Figure 2A,B). This realm covers 5 Baltimore classes, including all three classes of viruses with RNA genomes that replicate without a DNA intermediate as well as viruses with RNA or DNA genomes whose replication involves a reverse transcription stage. The study of the large-scale evolution of this virus realm is, out of necessity, based on comparison and phylogenetic analysis of RdRP/RT, the only VHG that is conserved across the realm. Notably, no other genes come even close to RdRP/RT with respect to the conservation among ribovirians (ending “virians” refers to all members of a realm (Koonin, Kuhn, et al., 2024)): the most common of these, such as the single jelly-roll (SJR) capsid proteins (Figure 2C), the principal building block of icosahedral ribovirion virions, are shared only by subsets of the viruses in the realm (Wolf, Kazlauskas, et al., 2018). Thus, comparative analysis of these genes can shed light on the evolution of particular groups of ribovirians but not of the entire realm. Indeed, virions of ribovirians are highly diverse with respect to morphology and complexity, and are built from many structurally unrelated major capsid proteins (Figure 2C,D). Thus, if virion morphogenesis proteins of ribovirians were chosen for realm definition, over a dozen realms would have to be created for the classification of all RNA and RT viruses.

The realm *Riboviria* is currently divided into two kingdoms, *Orthornavirae*, viruses that encode RdRP

and have no DNA stage in their replication cycles, and *Pararnavirae*, viruses that encode RTs and replicate via alternating RNA and DNA stages (Figure 2A). The two kingdoms share no conserved genes other than RdRP and RT. All RdRPs and RTs share a common structural fold (Figure 2B), and their sequences can be confidently aligned to produce a single phylogenetic tree that is not fully resolved but does contain major, robust clades (see below). Moreover, among the polymerases sharing a conserved core fold, known as the Palm domain or RNA recognition motif (RRM), RdRPs and RTs are relatively close, forming a single clade (Xiong and Eickbush, 1990). Thus, beyond doubt, all RdRPs and RTs evolved from a single ancestral enzyme that possessed polymerase activity, either RdRP or RT, or both. However, this homology does not actually mean that the two kingdoms share a common viral ancestry. All known RdRP-encoding replicators are viruses or derivatives of viruses that have lost genes for virion proteins (see below). Most likely, the common ancestor of orthornaviraens (ending “viraens” refers to all members of a kingdom (Koonin, Kuhn, et al., 2024)) was a simple virus encoding the RdRP and a SJR capsid protein. Orthornaviraens are ubiquitous in eukaryotes, and as recently demonstrated by metatranscriptome analysis, are also common in bacteria (Neri et al., 2022; Edgar et al., 2022; Zayed et al., 2022; Hou et al., 2024). Thus, the ancestral orthornaviraens most likely infected bacterial hosts, having emerged at an early stage of life evolution. Notably, despite considerable effort, orthornaviraens and more generally ribovirians have not been thus far detected in archaea, the causes of this potential exclusion remaining unclear.

Pararnaviraens have a completely different provenance, apparently, descending from group II introns, which are, in effect, prokaryotic retrotransposons (Gladyshev and Arkhipova, 2011; Toro and Nisa-Martinez, 2014). Pararnaviraens evolved only in eukaryotes as a result of the capture of the (nucleo)capsid proteins (Krupovic, Blomberg, et al., 2018). Thus, there is no evidence of common origin of orthornaviraens and pararnaviraens from a viral ancestor. Furthermore, there are strong indications that two orders within *Pararnavirae*, *Ortervirales* and *Blubervirales*, have evolved from distinct families of transposons by recruiting structurally unrelated major virion proteins (Krupovic, Blomberg, et al., 2018; Gong and G. Z. Han, 2018; Gong and G. Z. Han, 2025)

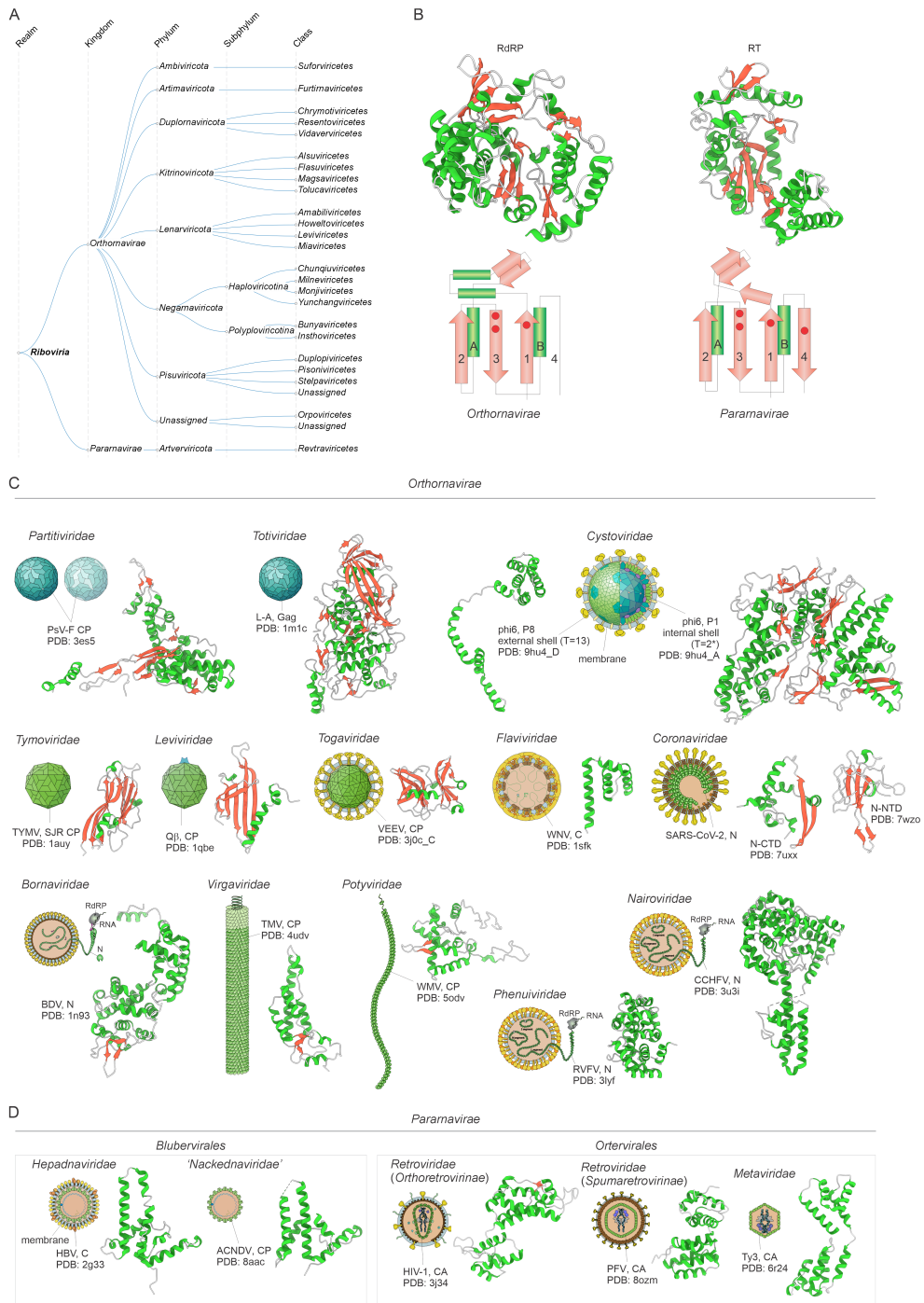


Figure 2. Caption continued on next page.

Figure 2. (cont.) Megataxonomy of viruses: realm *Riboviria*. (A) Taxonomic structure: from realm down to classes. The taxonomic tree was retrieved from the ICTV website (Black et al., 2026). (B) Structural models and topology diagrams of the hallmark proteins of *Riboviria*, RNA-dependent RNA polymerase (RdRP; PDB: 1ra7) and reverse transcriptase (RT; PDB: 7o0h). (C,D) Virion architectures and structural models of major capsid proteins (CP or CA) and nucleocapsid (N) proteins in the kingdoms *Orthornavirae* (C) and *Pararnavirae* (D). The proteins were selected to represent the diversity of structural folds in *Riboviria*. Virion diagrams were obtained from ViralZone (De Castro et al., 2024). The structures are denoted with the corresponding PDB accession numbers. Abbreviations: SJR CP, single jelly-roll capsid protein; C, core protein; PsV-F, *Penicillium stoloniferum* virus F; TYMV, turnip yellow mosaic virus; VEEV, Venezuelan equine encephalitis virus; WNV, West Nile virus; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; BDV, Borna disease virus; TMV, tobacco mosaic virus; WMV, watermelon mosaic virus; RVFV, Rift Valley fever virus; CCHFV, Crimean-Congo hemorrhagic fever virus; HBV, hepatitis B virus; ACNDV, African cichlid nakednavirus; HIV-1, human immunodeficiency virus 1; PFV, prototype foamy virus.

(Figure 2D). Accordingly, following the principle of viral realm monophyly, the current realm *Riboviria* should be split into at least three separate realms, one including all orthornaviraens and two more for the two independently evolved branches of pararnaviraens.

Initial phylogenetic analysis of orthornaviral RdRPs produced a tree with 5 major clades that were designated phyla (Wolf, Kazlauskas, et al., 2018; Wolf, Silas, et al., 2020). Three of the 5 phyla consist of positive-sense RNA viruses (although one phylum, *Pisuviricota*, also includes some groups of dsRNA viruses), one of negative-sense RNA viruses, and the last one of dsRNA viruses. The monophyly of each phylum is strongly statistically supported but the relationships among them remain largely uncertain (Neri et al., 2022; Edgar et al., 2022; Zayed et al., 2022). However, the basal position of phylum *Lenarviricota*, which consists of riboviruses infecting bacteria and their direct descendants infecting eukaryotes, is robust in RdRP trees rooted by RT, implying that all the diverse orthornaviraens of eukaryotes evolved from bacterial ancestors.

Metatranscriptome mining led to a manyfold increase in the known diversity of orthornaviraens, including several branches without a close relationship to any of the 5 large phyla that are candidates for new phyla (Neri et al., 2022; Edgar et al., 2022; Zayed et al., 2022; Hou et al., 2024). Two of these that have already been approved by the ICTV are of particular note. The new phylum *Artimaviricota* includes viruses with genomes consisting of two segments of dsRNA, one of which encodes a distinct polymerase that, as suggested by structural comparison, could be an evolutionary intermediate between RdRP and RT (Urayama, Fukudome, Hirai, et al., 2024). Artimaviricots (ending “viricots” refers to all members of a phylum (Koonin, Kuhn, et al., 2024)) were discovered in metatranscriptomes from hot springs dominated

by hyperthermophilic bacteria, which are the likely hosts of these viruses. Prokaryotic hosts are suggested also by the genome organization of artimaviricots, with genomic segments encompassing multiple protein-coding genes preceded by ribosome-binding sites (RBS or Shine-Dalgarno sequences).

More generally, perhaps, the greatest insight from the metatranscriptome mining for ribovirians is the dramatic expansion of the number and diversity of orthornaviraens associated with bacterial hosts, even if in many cases tentative. Apart from the discovery of an unexpected plentitude of previously unknown lenarviricots, multicistronic genome organization, with the telltale RBS preceding each protein-coding region, in addition to artimaviricots, was discovered in picobirnaviruses, paraxenoviruses and several branches of partitiviruses, all viruses with dsRNA genomes in phylum *Pisuviricota* (Neri et al., 2022; Krishnamurthy and D. Wang, 2018; D. Wang, 2022; Yoshida et al., 2025). Furthermore, it has been shown that many picobirnaviruses, initially thought to infect eukaryotes, encode functional bacterial lysins, strongly supporting the bacterial host for these viruses (Gan and D. Wang, 2023). In addition, several deep branches, representatives of putative new phyla, showed the same features suggestive of prokaryotic hosts. Thus, metatranscriptomics shattered the traditional view that riboviruses were primarily characteristic of eukaryotes, showing that RNA bacteriophages are a substantial, diverse component of prokaryotic viromes that has been largely overlooked before large metatranscriptome sequence datasets became available. Moreover, the mixing of viruses with known eukaryotic hosts with those predicted to infect bacteria among the partitiviruses suggests that, in some groups of orthornaviraens, particularly those that, like partitiviruses, harbor two or more genomic segments, each carrying a single gene, the switch from prokaryotic to

eukaryotic hosts or vice versa is comparatively easy and occurred on multiple occasions (Neri et al., 2022; Urayama, Fukudome, Hirai, et al., 2024). The remarkable ability of partitiviruses to replicate in highly diverse hosts has been demonstrated experimentally in eukaryotes, where the same virus has been shown to replicate in hosts from three eukaryotic kingdoms, including fungi, plants and animals (Telengech et al., 2024). How such simple viruses can overcome the physical barriers of the host envelope and defense systems in highly distinct cell types is an intriguing question.

Perhaps, an even more startling discovery is the 7th recognized phylum of orthornaviraens, *Ambiviricota*, which includes fungal parasites with covalently closed circular (ccc) RNA genomes encoding a distinct RdRP without clear affinity with any other viruses and an uncharacterized protein (Kuhn, Botella, et al., 2024). Ambivirus genomes resemble those of viroids in that both are circular and contain ribozymes, RNA segments with distinct secondary structures endowed with catalytic activity (B. D. Lee et al., 2023; Forgia et al., 2023). The mechanistic interplay between the ribozymes and RdRP during the ambiviral replication cycle remains to be elucidated. Ambiviricots combine features of two virus realms, *Riboviria* and *Ribozyviria* (see below), but given the presence of the RdRP, the hallmark of *Orthornavirae* within *Riboviria*, the realm and kingdom are not in doubt.

Mapping the gene content and genome organization of orthornaviraens to the RdRP tree allowed an approximate reconstruction of evolutionary trajectories, revealing the major trend of genome growth and complexification via accretion of genes captured from hosts and other viruses. The ancestral orthornaviraen is inferred to have had a small genome of perhaps about 3 kb, with two genes only, encoding the RdRP and a capsid protein, most likely, with the SJR fold (Wolf, Kazlauskas, et al., 2018). Subsequent evolution involved a substantial increase in genome size occurring independently in different lineages, up to 64 kb in some nidoviruses (Neuman et al., 2025; Zhong et al., 2025; Dolja, 2025). Acquisitions that occurred in parallel in different lines of descent included RNA helicases that apparently enabled efficient replication of larger RNA genomes, proteases that cleave viral polyproteins into mature, functional proteins and a variety of genes involved in virus–host

interaction, particularly inhibition of host defenses and movement proteins in plant viruses.

Pararnaviraens are far less diverse than orthornaviraens, with only one phylum and class, which splits into two orders, *Ortervirales* and *Blubervirales* (Krupovic, Blomberg, et al., 2018) (Figure 2A). Ortervirals (ending “virals” refers to all members of an order (Koonin, Kuhn, et al., 2024)) include the numerous animal retroviruses with RNA genomes replicating via a DNA intermediate along with plant caulimovirids (ending “virids” refers to all members of a family (ibid.)) with dsDNA genomes replicating via an RNA intermediate and three families of viruses that are often considered as “long terminal repeat (LTR) retrotransposons”, namely, *Metaviridae*, *Pseudoviridae* and *Belpaoviridae*. Members of the three families share with the animal retrovirids homologous RTs, aspartic proteases and a set of structural proteins, namely, capsid and nucleocapsid proteins, that form virus particles (Krupovic and Koonin, 2017a) (Figure 2D). Thus, according to the virus definition given above, “LTR retrotransposons” are bona fide viruses. Recently, a fourth potential family of ortervirals, described as Troyka “retrotransposons”, has been reported (Kojima, 2025), and additional divergent lineages of RT viruses are likely to be discovered in the near future.

Blubervirales includes human hepatitis B virus and its relatives infecting other vertebrate animals as well as a diverse group of related non-enveloped viruses informally known as nakednaviruses (Figure 2D). More recently, the reach of blubervirals was extended to rotifers, a group of microscopic and near-microscopic pseudocoelomate animals, by a family-level lineage, dubbed proto-nakednaviruses, which in the RT phylogeny form a basal group to hepadnavirids and nakednaviruses (Gong and G. Z. Han, 2025). Blubervirals possess DNA genomes that replicate via an RNA intermediate and their structural proteins are unrelated to those of ortervirals (Wynne et al., 1999) (Figure 2D).

Pararnaviraens are limited in their spread to eukaryotic hosts and apparently evolved from non-viral retroelements, such as group II self-splicing introns and retrotransposons, by acquiring host proteins and exapting them for structural roles in the virions (Krupovic and Koonin, 2017b). Furthermore, the two orders within this kingdom apparently evolved at two independent points of origin from different

retrotransposon families (Gong and G. Z. Han, 2018; D. Wang, 2022; Malik and Eickbush, 2001). Whereas the exact ancestor of ortervirals remains obscure, blubervirals have apparently evolved from a distinct group of retroelements known as *HEART* (Gong and G. Z. Han, 2018).

4.3. Four realms of viruses with ssDNA genomes

Until recently, the vast majority of viruses with ssDNA genomes and two groups of viruses with small dsDNA genomes (papillomavirids and polyomavirids) were classified in a single realm, *Monodnaviria*. The realm was held together by a single VHG that encodes a distinct endonuclease of the HUH superfamily (or its inactivated derivative) involved in the initiation of the genome replication via the rolling circle (or rolling hairpin) mechanism (Ilyina and Koonin, 1992; Chandler et al., 2013). However, recently, upon considerable expansion of the ssDNA virome, largely through metagenomics, and improved understanding of evolutionary relationships among these viruses, the realm was split into four monophyletic realms, corresponding to the four initially defined kingdoms (Figure 3A).

Realm *Efunaviria*, with the kingdom *Loebvirae*, consists of filamentous prokaryotic viruses with circular ssDNA genomes, including some well-characterized phages of family *Inoviridae* such as M13, f1 and fd (Roux, Krupovic, Daly, et al., 2019). The morphogenetic module and the overall virion assembly mechanism of efunavirians are conserved. All efunavirians encode a highly hydrophobic capsid protein consisting of a single α -helix, which coats the circular ssDNA genome upon its extrusion through the cytoplasmic membrane with the aid of the conserved virus-encoded FtsK-family ATPase (Rakonjac et al., 2024) (Figure 3A). By contrast, the genome replication enzymes vary. Although some members of the realm employ HUH endonucleases, the majority relies on a variety of rolling circle replication initiation endonucleases of the Rep_trans superfamily that is not homologous to the HUH superfamily (Roux, Krupovic, Daly, et al., 2019), whereas others use transposases that generate ssDNA molecules as a transposition intermediate (Krupovic and Forterre, 2015) (Figure 3B). Efunavirians infect widely diverse bacterial hosts, suggesting long-term association. However, putative efunavirians were also discovered

in some methanogenic archaea (Roux, Krupovic, Daly, et al., 2019), suggesting that the emergence of this virus realm could be more ancient, potentially dating back to the LUCA (Krupovic, Dolja, et al., 2020).

Realm *Volvereviria*, with the kingdom *Sangervirae*, represents a relatively uniform group of prokaryotic viruses, exemplified by the iconic bacteriophage phiX174 and other phages of the recently created class *Microviricetes* (formerly, family *Microviridae*). Volverevirians form small icosahedral capsids built of a distinct SJR-fold capsid protein (Dokland et al., 1997; H. Lee et al., 2022; Mietzsch, Kailasan, et al., 2024; Bardy et al., 2025) (Figure 3A), which is commonly used as a marker for the discovery and classification of volverevirians (Roux, Krupovic, Poulet, et al., 2012; Kirchberger, Martinez and Ochman, 2022; Olo Ndela et al., 2023; Lund et al., 2024). These viruses have circular ssDNA genomes replicated by the rolling circle mechanism and uniformly (thus far) encode the HUH endonuclease (Doore and Fane, 2016; Kirchberger and Ochman, 2023) (Figure 3B). Volverevirians can be either strictly lytic, typically lysing the host by inhibiting peptidoglycan synthesis (Chamakura and Young, 2020), or temperate, integrating into the bacterial genome using the host recombination machinery (Krupovic and Forterre, 2011; Kirchberger and Ochman, 2020). All experimentally characterized members of this realm are phages, but for the majority of volverevirians identified via metagenomics, the hosts remain unknown, so it cannot be excluded that some of them replicate in archaea.

Realm *Floreoviria*, with the kingdom *Shotokuvirae*, encompasses a highly diverse group of eukaryotic viruses with small ssDNA or dsDNA genomes, which in most members are circular, but can also be linear (e.g., families *Parvoviridae*, *Bidnaviridae*, *Oomyviridae*). Florevirians form icosahedral capsids from SJR capsid proteins (Mietzsch, Bennett, et al., 2025) which, however, appear to have a distinct origin from the SJR capsid proteins of volverevirians (Krupovic and Koonin, 2017b). Furthermore, in some lineages (e.g., *Bacilladnaviridae*, *Naryaviridae*, *Nenyaviridae*, and cruciviruses), the ancestral SJR capsid protein has been replaced by distinct SJR capsid protein variants from RNA viruses, allowing for the formation of larger ($T = 3$) capsids (Roux, Enault, et al., 2013; Kazlauskas, Dayaram, et al., 2017;

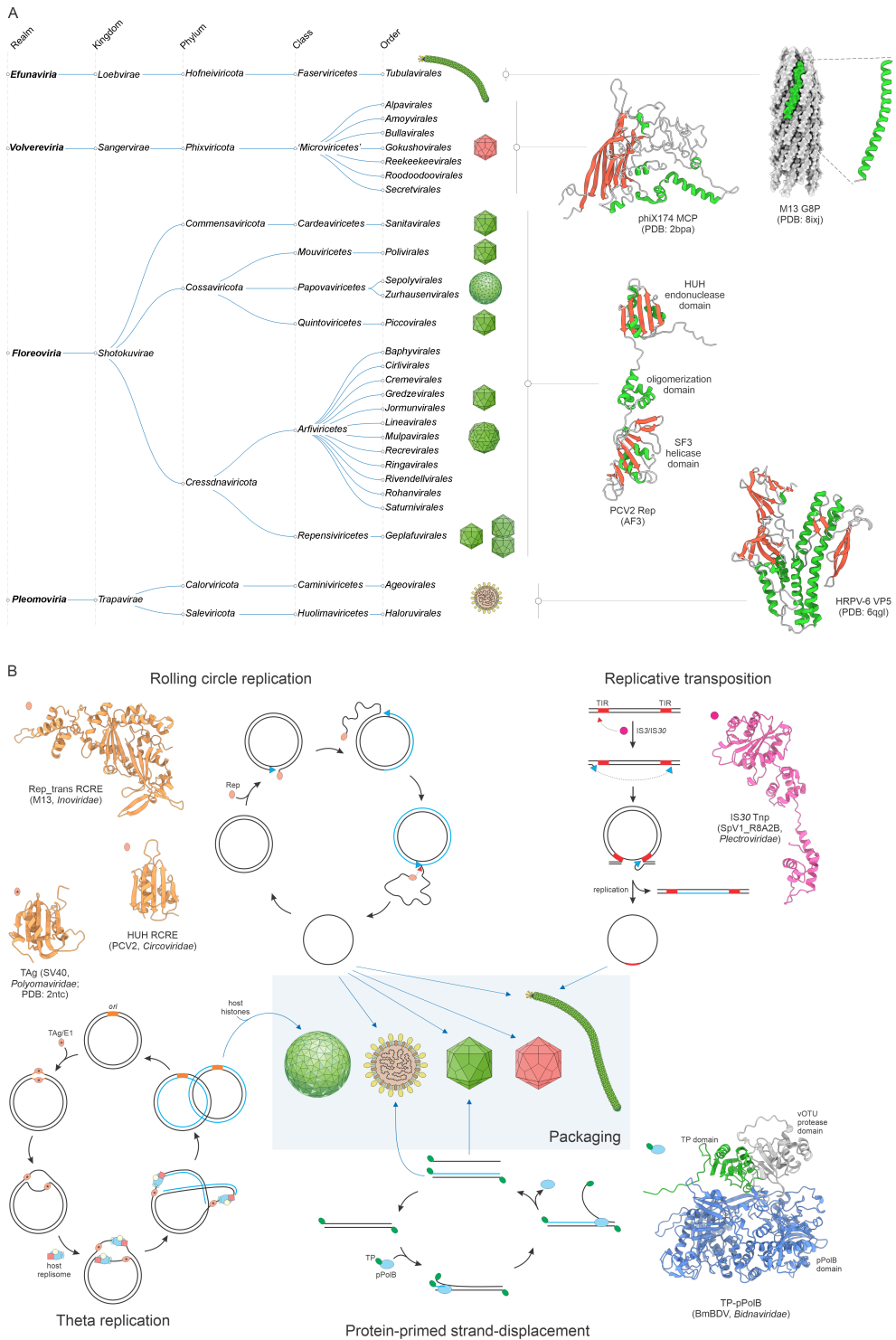


Figure 3. Caption continued on next page.

Figure 3. (cont.) Megataxonomy of viruses: the four realms of viruses with small ssDNA or dsDNA genomes. (A) Taxonomic structure, from realms down to orders, and structural models of the realm-specific hallmark proteins: major capsid proteins (*Efunaviria* and *Volvereviria*), rolling circle replication initiator containing the HUH superfamily endonuclease and superfamily 3 helicase domains (*Floreoviria*), and membrane fusion protein VP5 (*Pleomoviria*). The structures are denoted with the corresponding PDB accession numbers. AF3 denotes structures modeled with AlphaFold3. The taxonomic tree was retrieved from the ICTV website (Black et al., 2026) and modified. (B) The three distinct replication mechanisms catalyzed by unrelated proteins and yielding ssDNA genomes, which are encapsidated into virus particles. Note that some members of the realm *Pleomoviria* encapsidate either dsDNA or ssDNA replicative intermediates. Also shown is the theta mode of genome replication initiated by the inactivated HUH endonuclease—superfamily 3 helicase of polyomavirids and papillomavirids. Abbreviations: BmBDV, Bombyx mori bidensovirus; TP, terminal protein; pPolB, protein-primed family B DNA polymerase; TAg, large T antigen; SV40, simian virus 40; ori, origin of replication; TIR, terminal inverted repeats; Tnp, transposase; PCV2, porcine circovirus 2; MCP, major capsid protein; HRPV-6, Halorubrum pleomorphic virus 6. Virion diagrams were obtained from ViralZone (De Castro et al., 2024).

de la Higuera et al., 2020; Krupovic and Varsani, 2022; Gebhard et al., 2025; Munke et al., 2022). A characteristic feature of viruses in this realm is the two-domain replication protein, consisting of the N-terminal HUH endonuclease and C-terminal superfamily 3 helicase domains (Kazlauskas, Varsani, et al., 2019) (Figure 3A). The ancestor of florevirians, in all likelihood, encoded the two-domain HUH superfamily endonuclease and replicated via the rolling circle mechanism, which is employed by the majority of the members of this realm (Figure 3B). However, in some descendants, this gene has been either inactivated (polyomavirids and papillomavirids), concomitant with a switch from the rolling circle to the theta-like replication mechanism (Iyer, Koonin, et al., 2005), or replaced by a protein-primed family B DNA polymerase (bidnavirids) (Krupovic and Koonin, 2014) (Figure 3B). An even more dramatic departure from the canonical gene combination is observed in members of the family *Anelloviridae* (phylum *Commensaviricota*, order *Sanitavirales*). Anellovirids retain the SJR capsid protein, albeit with major modifications not observed in other florevirians (Butkovic, Kraberger, et al., 2023; Liou et al., 2024; De Koch et al., 2025), but the ancestral replication initiation gene has been lost, with the genome replication apparently fully depending on the host replication machinery (Boisvert et al., 2025). Although the exact mechanism remains to be elucidated, it has been suggested that anellovirids employ recombination-dependent replication by recruiting the host DNA polymerase alpha and BTR (Bloom's syndrome helicase (BLM), topoisomerase III α , RMI1, and RMI2) complexes (ibid.), with the circular ssDNA genomes being produced by a process that resembles the formation of extratelomeric C-circles (J. Lee, J. Lee, et al., 2024; J. Lee, Sohn, et al., 2025).

Finally, the realm *Pleomoviria*, with the kingdom *Trapavirae*, includes several families of archaeal viruses with enveloped pleomorphic virions (Figure 3A). The virions resemble membrane vesicles with two major structural proteins, a spike protein with a unique structural fold responsible for host recognition and membrane fusion, and a matrix protein embedded in the viral envelope (El Omari et al., 2019; Y. Liu, Dyall-Smith, et al., 2022; Bignon et al., 2022). The membrane fusion protein is a signature of pleomovirians (Figure 3A). Initially, *Trapavirae* included a single family, *Pleolipoviridae*, of viruses infecting extremely halophilic archaea (Y. Liu, Dyall-Smith, et al., 2022; Alarcon-Schumacher et al., 2023). However, more recently, related viruses were also identified in hyperthermophilic, methanogenic and nano-sized hyperhalophilic archaea (Medvedeva, Borrel, et al., 2023; Baquero, Bignon, et al., 2024; Zhou et al., 2025), expanding the genetic diversity and host range of this virus group. The genomes of pleomovirians can be ssDNA or dsDNA, circular or linear. Accordingly, the replication mechanisms and the corresponding enzymes are highly variable, including several non-orthologous and even non-homologous rolling circle replication endonucleases (Kazlauskas, Varsani, et al., 2019; Baquero, Bignon, et al., 2024; Roine et al., 2010; Y. Wang et al., 2026), primases (Alarcon-Schumacher et al., 2023) or protein-primed family B DNA polymerases (Y. Liu, Dyall-Smith, et al., 2022) (Figure 3B). Thus, it is the shared morphogenetic rather than genome replication module that holds this realm together.

Viruses of the four realms, *Efunaviria*, *Volvereviria*, *Floreoviria* and *Pleomoviria*, formerly joined under *Monodnaviria* typically have small genomes (2–10 kb; Figure 1C) and their evolution appears to be tightly intertwined with that of small plasmids, largely replicating via the rolling circle mechanism,

and apparently, ancestors of the four realms evolved from different families of such plasmids (Kazlauskas, Varsani, et al., 2019). In all cases of independent origins of viruses with small ssDNA genomes, the capsid protein genes were captured by the emerging viruses independently from other viruses, such as orthornaviraens, or from host genes.

4.4. Realms *Varidnaviria* and *Singelaviria*: the epitome of viral diversity

Realm *Varidnaviria* consists of an enormous diversity of dsDNA viruses infecting bacteria, archaea, and eukaryotes (Figure 4A) that typically have an icosahedral capsid built of double jelly-roll (DJR) major capsid protein (MCP), the principal VHG that holds the realm together (Krupovic and Bamford, 2008), and SJR minor capsid protein (penton), which form hexagonal and pentagonal capsomers, respectively (N. G. Abrescia et al., 2004) (Figure 4B,C). Most varidnavirians also share another VHG encoding a genome packaging ATPase of the FtsK-HerA superfamily (Iyer, Makarova, et al., 2004). The varidnavirians display a remarkable variation in virion sizes, morphologies and complexity. Most icosahedral virions in this realm contain an internal lipid membrane, sandwiched between the protein capsid and the genome, which likely represents the ancestral trait of *Varidnaviria* (Figure 4D). However, in some lineages (e.g., adenovirids), the internal membrane was lost, whereas in others, an additional external membrane was added (e.g., iridovirids), and/or a second internal icosahedral shell built from unrelated capsid proteins has been added (e.g., asfarvirids and faustoviruses) (Figure 4D). Even more dramatic deviations from the canonical structural layout occurred in several groups of varidnavirians, e.g., pandoraviruses and pithoviruses, where the DJR MCP was replaced by unrelated proteins, resulting in odd capsid shapes (Philippe et al., 2013; Legendre, Bartoli, et al., 2014; Souza et al., 2019; Krupovic, Yutin, et al., 2020; Queiroz et al., 2023) (Figure 4D). Other groups of viruses within *Varidnaviria*, such as poxvirids, encode a homolog of the DJR MCP but incorporate it only into intermediates of the virion morphogenesis, whereas the odd-shaped mature capsids consist of unrelated viral proteins (Coulilbaly, 2024).

Varidnaviria consists of two kingdoms, *Bamfordvirae* and *Abadenavirae* (Koonin, Fischer, et al.,

2024). The kingdom *Abadenavirae* includes all bacterial and archaeal viruses with DJR MCP and relatively small dsDNA or ssDNA genomes (e.g., finnlakevirids), except for one family of bacterial viruses, *Tectiviridae*. The kingdom *Bamfordvirae* includes tectivirids and a broad diversity of eukaryotic viruses, characterized by highly diverse genome and virion sizes, including megagenomoviruses with genomes in excess of 4 Mb (Buscaglia et al., 2025) (Figure 4A).

The evolutionary scenario for *Varidnaviria* derives the entire, striking diversity of eukaryotic varidnavirians from a tectivirid ancestor (Figure 4D), possibly, passed to the emerging eukaryote by the mitochondrial endosymbiont (Krupovic and Koonin, 2015; Krupovic, Kuhn, Fischer, et al., 2024). Notably, members of a diverse assemblage of bamfordviraens with relatively small genomes (15–35 kb) were originally recognized as DNA transposons known as polintons (after Polymerase and Integrase, two proteins encoded in their genomes; alternatively, called mavericks) that are integrated in genomes of a broad variety of eukaryotes, in some cases, in numerous copies (Kapitonov and Jurka, 2006; Pritham et al., 2007; Barreat and Katzourakis, 2021; Jeong et al., 2023). Subsequent analyses have shown, however, that polintons encode a complete morphogenetic module comprising typical DJR-MCP, penton protein with the SJR fold, capsid maturation protease and the packaging ATPase, strongly suggesting that polintons are bona fide varidnavirians (Krupovic, Bamford and Koonin, 2014). Currently, polintons and numerous polinton-like viruses, of which many but not all encode integrases and are known or predicted to integrate into the host genomes, are classified into subphylum *Polisuviricotina* of phylum *Preplasmiviricota* that also includes the tectivirids (Koonin, Fischer, et al., 2024). Many polisuviricotins are virophages, that is, symbionts (parasites or commensals) of large viruses of phylum *Nucleocytoviricota* (La Scola et al., 2008; Fischer, 2021; Roux, Fischer, et al., 2023; Roitman et al., 2023).

Similar to viruses with smaller DNA genomes, varidnavirians employ a variety of genome replication strategies and, accordingly, encode a range of non-homologous replication proteins, ranging from Rep_trans and HUH endonucleases for rolling circle replication (e.g., finnlakevirids and corticovirids, respectively (Laanto et al., 2017; Krupovic and Bamford, 2007)) to protein-primed family B DNA

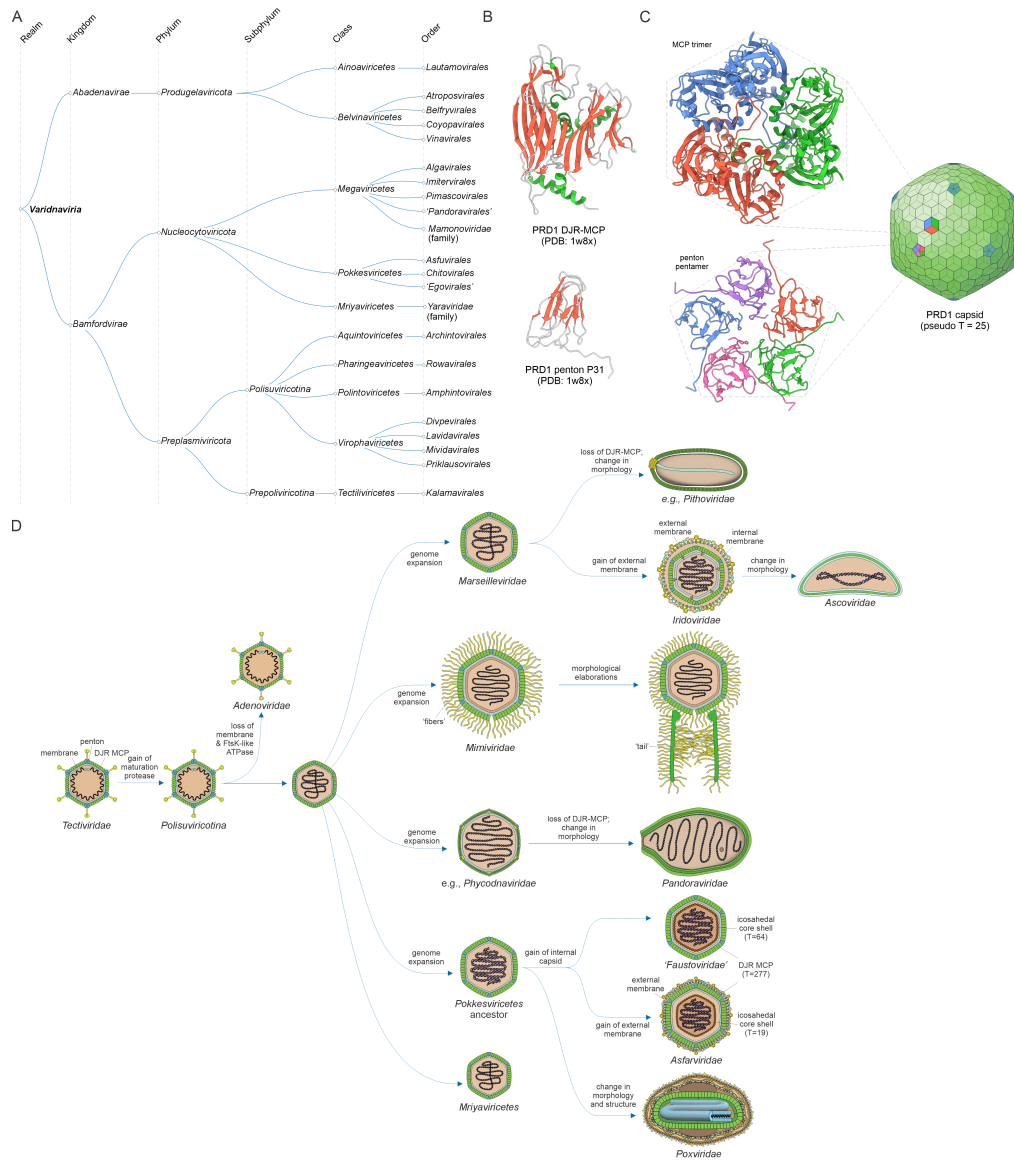


Figure 4. Megataxonomy of viruses: realm *Varidnaviria*. (A) Taxonomic structure, from realms down to orders. The taxonomic tree was retrieved from the ICTV website (Black et al., 2026). (B) Structural models of the double jelly-roll major capsid protein (DJR-MCP; top) and single jelly-roll penton protein (bottom) of bacteriophage PRD1 (*Tectiviridae*). The structures are denoted with the corresponding PDB accession numbers. (C) Structural models of the MCP trimer and penton pentamer, which form hexagonal and pentagonal capsomers. In tectivirids and adenovirids, the two types of capsomers are arrayed to form icosahedral capsids with a pseudo $T = 25$. (D) Diversity of virion organizations and complexity among varidnavirians and proposed evolutionary scenarios. Virion diagrams were obtained from ViralZone (De Castro et al., 2024).

polymerases and their various derivatives (e.g., tectivirids and adenovirids (Krupovic, Kuhn, Fischer, et al., 2024)) to family A DNA polymerases (e.g., sputniviroid (Iyer, Abhiman, et al., 2008)) to the complete or near-complete replisome in most

nucleocytoviricots (Koonin and Yutin, 2019; Koonin and Yutin, 2018). Notably, major variations are observed within *Nucleocytoviricota*: members of the class *Mriyaviricetes*, the nucleocytoviricots with the smallest genomes that apparently comprise the basal

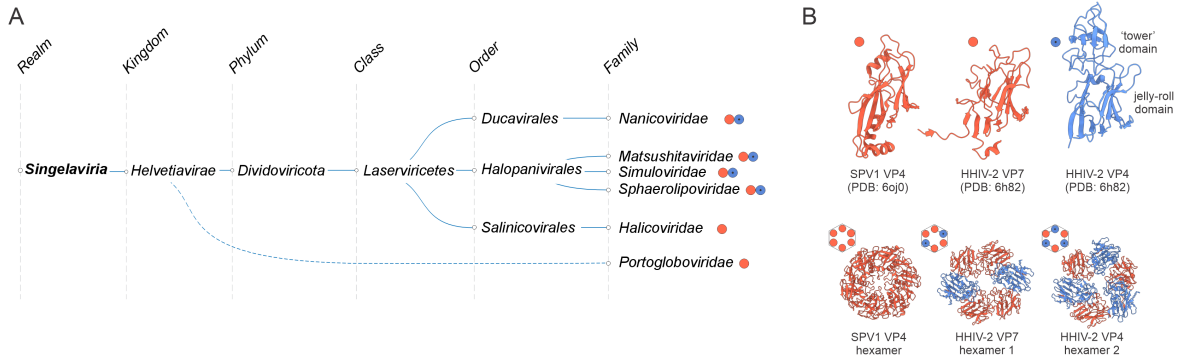


Figure 5. Megataxonomy of viruses: realm *Singelaviria*. (A) Taxonomic structure, from realms down to families. Family *Portogloboviridae* is connected to the realm by a dashed line because it is currently not formally included in *Singelaviria*, but is likely to represent the ancestral state of the realm. (B) Structural models of the single jelly-roll fold major capsid proteins (MCP) and their oligomers. Some members of the realm, e.g., *Sulfolobus polyhedral virus 1* (SPV1; *Portogloboviridae*), encode a single MCP VP4 (red) which forms homohexameric capsomers, whereas others, e.g., *Haloarcula hispanica* icosahedral virus 2 (HHIV-2; *Sphaerolipoviridae*), encode two paralogous MCPs, VP7 (red) and VP4 (blue), which form two types of heterohexameric capsomers, hexamers 1 and 2, respectively. The structures are denoted with the corresponding PDB accession numbers.

branch of this phylum, encode an HUH endonuclease but no DNA polymerase and are predicted to replicate via the rolling circle mechanism (Yutin, Mutz, et al., 2024). Even relatively closely related viruses (e.g., those from the same order) may encode unrelated replication modules. Furthermore, the presence of the conserved HUH endonuclease of *Mriyaviricetes* formally connects this viral class to two of the realms of viruses with ssDNA genome, for which this protein is a hallmark. However, the overall phylogenomic analysis leaves no doubt about the assignment of *mriyaviricetes* to *Nucleocyotviricota*.

Realm *Singelaviria* includes bacterial and archaeal viruses with linear or circular dsDNA genomes and icosahedral capsids built from SJR MCPs (Figure 5A,B). All structurally characterized, formally classified singelavirians encode two paralogous SJR MCPs, which form two types of heterohexameric capsomers (Figure 5B), resembling the homotrimeric capsomers of varidnavirians (Figure 4C). Furthermore, similar to most varidnavirians, the icosahedral virions of singelavirians contain an internal membrane (Demina et al., 2023). Thus, initially, it has been assumed that fusion of the two singelavirian SJR MCP genes yielded the DJR MCP (Rissanen et al., 2013; Santos-Perez et al., 2019; De Colibus et al., 2019). This scenario was further supported by the fact that, like varidnavirians, singelavirians encode an FtsK-HerA superfamily genome packaging ATPase (Strömsten et al., 2005). Accordingly, singelavirians

were included in the realm *Varidnaviria* as kingdom *Helvetiavirae*. However, more recently, a detailed comparison of protein structures showed that MCPs of singelavirians and varidnavirians evolved from distinct cellular proteins, an SJR carbohydrate-binding protein and DJR glycoside hydrolase, respectively (Krupovic, Makarova, et al., 2022). Accordingly, kingdom *Helvetiavirae* has been reclassified as a separate realm.

Further exploration of virus diversity led to the identification of relatively close singelavirian relatives that encode the FtsK-HerA genome packaging ATPase and a single SJR MCP (e.g., halicovirids), implying homohexameric capsomers (Zhou et al., 2025; Makarova et al., 2014). Finally, archaeal viruses of the *Portogloboviridae* family (Y. Liu, Ishino, et al., 2017; Prangishvili et al., 2021), currently not formally assigned to *Singelaviria* (Figure 5A), might even more closely resemble the ancestral state. *Portoglobovirids* lack the genome packaging ATPase and encode a single SJR MCP, which closely resembles MCPs of singelavirians and forms homohexamers (Figure 5B), as predicted for halicovirids (F. Wang, Y. Liu, et al., 2019).

4.5. Realm *Duplodnaviria*, the champions of the viral world

Realm *Duplodnaviria* includes the most abundant viruses on Earth, namely, the tailed dsDNA bacterial

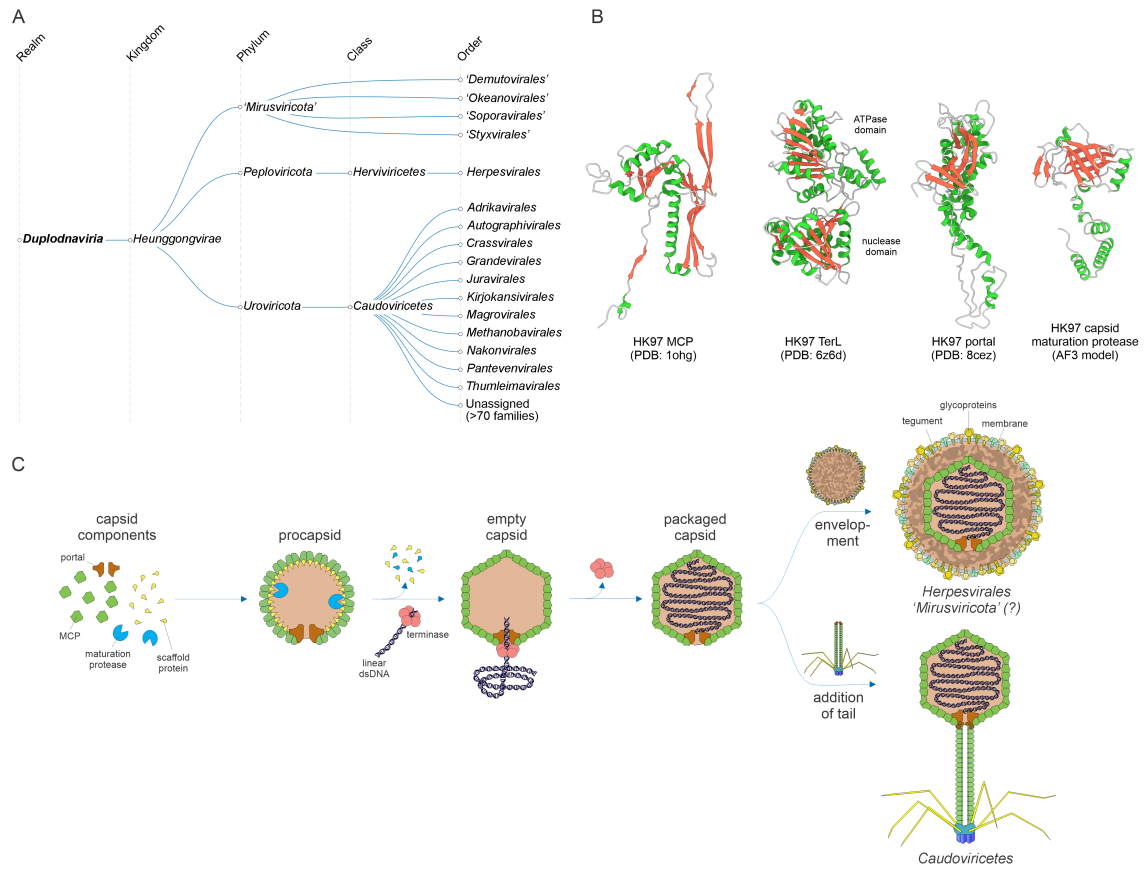


Figure 6. Megataxonomy of viruses: realm *Duplodnaviria*. (A) Taxonomic structure: from realm down to orders. The taxonomic tree was retrieved from the ICTV website (Black et al., 2026) and modified to include the proposed phylum “*Mirusviricota*”. (B) Structural models of the hallmark proteins comprising the morphogenetic module conserved in duplodnavirians, exemplified by the major capsid protein (MCP), large subunit of the terminase (TerL), portal protein, and capsid maturation protease of bacteriophage HK97. (C) Virion morphogenesis in the three phyla of duplodnavirians. The assembly pathway was depicted using the modified virion diagrams obtained from ViralZone (De Castro et al., 2024).

and archaeal viruses (class *Caudoviricetes*) (Dion et al., 2020; Cai et al., 2021; Adriaenssens et al., 2025), the recently discovered “mirusviruses” associated with unicellular eukaryotic hosts (putative new phylum “*Mirusviricota*”) (Gaia et al., 2023), and animal viruses of the order *Herpesvirales* (Dotto-Maurel et al., 2025) (Figure 6A). All these viruses share a distinct morphogenetic module that consists of four VHGs encoding a HK97-fold MCP, a genome packaging ATPase-nuclease (known as large terminase subunit) that is distinct from the functional counterpart in varidnavirians, a portal protein, and a distinct capsid maturation protease (Iranzo, Krupovic, et al., 2016) (Figure 6B). The four VHGs play key roles in capsid assembly, maturation and

genome packaging (Figure 6C) and are highly conserved throughout *Duplodnaviria*. The assembly pathways following the genome packaging differ in a host-dependent manner. In prokaryote-infecting *Caudoviricetes*, the packaged capsid is combined with a preassembled tail (Dion et al., 2020; Baquero, Y. Liu, et al., 2020), whereas in eukaryote-infecting herpesvirals and mirusviruses, the capsid is enveloped with a lipid membrane (Crump, 2018; Chung et al., 2025) (Figure 6C).

Realm *Duplodnaviria* has a simple taxonomic structure including a single kingdom and three phyla (Figure 6A), *Uroviricota* (caudoviricetes), *Peploviricota* (herpesvirals) and the yet to be formally recognized “*Mirusviricota*” (mirusviricots).

Notwithstanding this uniformity at the top ranks of the taxonomy, the diversity of uroviricots is enormous, resulting in a constant flux of taxonomy reorganization (Adriaenssens et al., 2025; Y. Liu, Demina, et al., 2021; Turner et al., 2023).

The range of genome size and complexity of duplodnavirians nearly matches that of varidnavirians, spanning from about 10 kb to about 750 kb, all within *Uroviricota* (Luque et al., 2020; Al-Shayeb, Sachdeva, et al., 2020). Until the recent discovery of “*Mirusviricota*”, the evolutionary history of this realm appeared enigmatic, with a broad gulf separating tailed viruses of prokaryotes (bacteria and archaea) from herpesviruses that have been only identified in animals. However, “*Mirusviricota*”, a vast phylum of viruses apparently associated with a broad variety of unicellular eukaryotic hosts, fill this gap, in all likelihood, representing the original diversity of duplodnavirians in eukaryotes and the ancestors of herpesvirals (Gaia et al., 2023; Medvedeva, Guyet, et al., 2026).

As in the case of other virus groups discussed above, duplodnavirians display a remarkable diversity of genome replication strategies, especially among bacterial and archaeal caudoviricetes (ending “viricetes” refers to all members of a class (Koonin, Kuhn, et al., 2024)), which employ all conceivable DNA replication strategies and encode suites of replication factors ranging from single-protein initiators to complete replisomes (Weigel and H. Seitz, 2006; Krupovic and Bamford, 2010; Kazlauskas, Krupovic and Venclovas, 2016; Filosof et al., 2017). Notable variations also exist in the replication modules of eukaryotic duplodnavirians. For instance, in the order *Herpesvirales*, family B DNA polymerases of orthoherpesvirids are not orthologous to those of malacoherpesvirids and alloherpesvirids (Kazlauskas, Krupovic, Guglielmini, et al., 2020), whereas many mirusviricots lack DNA polymerases altogether (Medvedeva, Guyet, et al., 2026).

4.6. Small realms *Adnaviria* and *Ribozyviria*

Realm *Adnaviria* consists of archaeal viruses with rigid or flexible filamentous virions that can be either non-enveloped or contain an external lipid membrane (Krupovic, Kuhn, F. Wang, et al., 2021; Krupovic, F. Wang, et al., 2025) (Figure 7A). The remarkable feature of these viruses is that their linear

dsDNA genome is stored in the virions in the A conformation (Baquero, Y. Liu, et al., 2020; DiMaio et al., 2015). Adnavirians lack conserved VHGs although members of some families do encode them, e.g., HUH endonuclease that is conserved in rudivirids (Oke et al., 2011). However, among themselves, adnavirians share several conserved genes, most notably, the MCP with a unique α -helical fold (Kasson et al., 2017; Y. Liu, Osinski, et al., 2018; F. Wang, Baquero, L. C. Beltran, et al., 2020; F. Wang, Baquero, Su, et al., 2020), first described in the rudivirid SIRV2 (DiMaio et al., 2015). Although some adnavirians encode a single MCP, many encode two paralogous MCPs; these bind dsDNA as homodimers and heterodimers, respectively (Figure 7B). Adnavirians were initially detected in hyperthermophilic archaea of the phylum *Thermoproteota*, but subsequent metagenomics surveys showed that these viruses are also associated with hosts from the other archaeal phyla (Laso-Perez et al., 2023; Tamarit et al., 2022).

The small realm *Ribozyviria* (Hepojoki et al., 2021; Walker et al., 2022) currently includes a single family, *Kolmioviridae* (Kuhn, Babaian, et al., 2024), which encompasses human hepatitis D viruses 1–8 (genus *Deltavirus*) and their relatives discovered in other animals (Wille et al., 2018; Chang et al., 2019; Paraskevopoulou et al., 2020; Bergner et al., 2021; Perez-Vargas et al., 2021) (Figure 7C). Ribozyvirians are viroid-like covalently closed circular (ccc) RNA replicons encoding a nucleocapsid protein (delta antigen, DAg). Similar to viroids, these viruses hijack the cellular transcription machinery for their genome replication and depend on other viruses (bluberviral hepatitis B virus in the case of deltaviruses, arenavirids in case of daletviruses, unknown for most of the others) for the formation of infectious enveloped virions (Khalfi et al., 2023; Szivovics et al., 2020). DAg adopts a unique α -helical fold and forms octamers (Figure 7D) (Cornillez-Ty and Lazinski, 2003). The viral pseudo-dsRNA is thought to wrap around the DAg octamers that mimic nucleosomes and facilitate the recruitment of the host RNA polymerase II and the associated chromatin remodeling complexes for viral RNA replication (Abeywickrama-Samarakoon et al., 2020; Yang et al., 2025). DAg also induces packaging of the kolmiovirid ribonucleoprotein complex into progeny virions. Recent evidence suggests that this complex is directly incorporated into the enveloped virions of their helper viruses

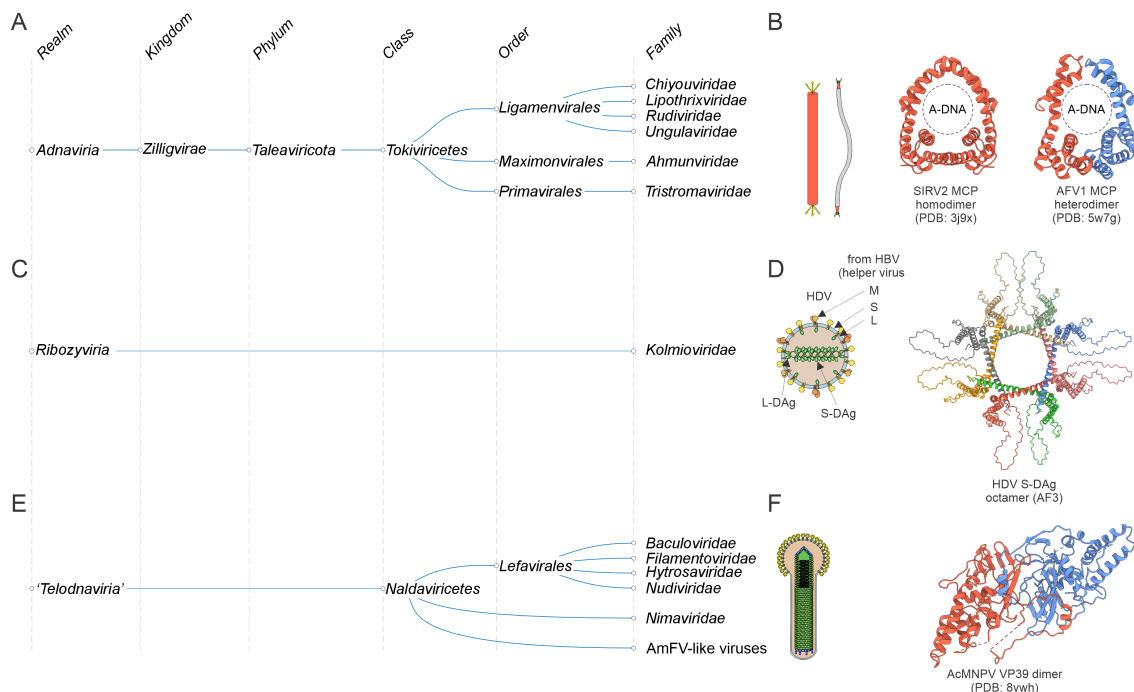


Figure 7. Megataxonomy of viruses: three small realms. (A) *Adnaviria*, taxonomic structure from realm down to families. (B) Virion architecture and structural model of the homodimeric and heterodimeric major capsid proteins of adnavirians, *Saccharolobus islandicus* rod-shaped virus 2 (SIRV2) and *Acidianus filamentous virus 1* (AFV1), respectively. The structures are denoted with the corresponding PDB accession numbers. (C) *Ribozoviria*, taxonomic structure from realm down to families. (D) Virion architecture and structural model of the nucleocapsid protein (Delta antigen) octamer of the prototype ribozovirion, hepatitis delta virus. The structural model was generated using AlphaFold3 (AF3). S-DAG and L-DAG, small and large delta antigens. S, M, L: hepatitis B virus glycoproteins. (E) “*Telodnaviria*”, taxonomic structure from realm down to families. “*Telodnaviria*” has not yet been formally recognized by the ICTV. (F) Virion architecture and structural model of the major capsid protein dimer of a well-characterized telodnavirion, *Autographa californica* multiple nucleopolyhedrosis virus (AcMNPV). The PDB accession number is indicated. AmFV, *Apis mellifera* filamentous virus.

(McKellar et al., 2026). No autonomous ribozovirians have been described to date, suggesting that all members of this realm are satellites of other viruses. Metagenome mining led to the discovery of a substantial variety of ribozy-like cccRNAs encoding distant homologs of the deltavirus nucleocapsid protein and most likely replicating in unicellular eukaryotes (Edgar et al., 2022; B. D. Lee et al., 2023). Thus, although *Ribozoviria* will probably remain a small viral realm, its true diversity appears to be much greater than reflected in the current taxonomy.

4.7. Are there more viral realms to be established?

The 10 realms described above encompass ~95% of the currently recognized virus families (Simmonds, Adriaenssens, Lefkowitz, et al., 2025). What about the

remaining 5% that are currently not assigned to any realm? It is highly likely that some of the unassigned families will be unified into additional realms. For instance, viruses with large dsDNA genomes currently classified into the class *Naldaviricetes* (van Oers et al., 2023; Guinet et al., 2024) were suggested to represent a separate realm, “*Telodnaviria*” (Johnstone et al., 2024) (Figure 7E). Viruses of this group, of which the best-known ones are baculoviruses due to their wide use as biocontrol agents in agriculture and as tools in molecular biology (Pidre et al., 2022), infect diverse arthropods, including insects and crustaceans, and form complex helical virions. The baculovirus MCP has a unique fold, unrelated to those of viruses from the other realms (Johnstone et al., 2024; Benning et al., 2024) (Figure 7F). Furthermore, members of the *Naldaviricetes* share a core gene set, including genes

involved in genome replication and transcription as well as per os infectivity factors (PIF) (van Oers et al., 2023), which are believed to form a distinct attachment and membrane fusion complex (Johnstone et al., 2024).

Another realm is likely to be created for archaeal viruses with dsDNA genomes and spindle-shaped or lemon-shaped virions. Spindle-shaped viruses are ubiquitous in archaea, associated with phylogenetically diverse hosts (Zhou et al., 2025; Krupovic, Quemin, et al., 2014). Due to their broad distribution in archaea, spindle-shaped viruses are thought to have been coevolving and diversifying with archaea ever since the emergence of the last archaeal common ancestor from the LUCA (Krupovic, Dolja, et al., 2020). Spindle-shaped viruses encode a distinct MCP which adopts a simple α -helical hairpin fold (F. Wang, Cvirkaite-Krupovic, et al., 2022; Z. Han et al., 2022). The highly hydrophobic MCP subunits polymerize into a helical assembly, which has a larger diameter in the central region where the genome is located (F. Wang, Cvirkaite-Krupovic, et al., 2022). Spindle-shaped viruses are highly diverse in terms of gene content and genome replication strategies. Some have linear genomes replicated by virus-encoded protein-primed family B DNA polymerase (Tamarit et al., 2022; Bath et al., 2006; Kim et al., 2019; Medvedeva, Sun, et al., 2022), whereas others have circular genomes and encode either HUH superfamily endonucleases (Zhou et al., 2025; Medvedeva, Sun, et al., 2022) or homologs of various host replisome components, such as RNA-primed family B DNA polymerase (Laso-Perez et al., 2023) or MCM helicase (Gorlas et al., 2012), or lack recognizable replication proteins. Due to their unique architecture and the fact that spindle-shaped viruses do not share conserved VHGs with other virus realms, they are likely to have an independent origin and thus qualify as a separate realm.

Several additional groups of archaeal viruses with enigmatic gene contents and unusual virion morphologies, such as bottle-shaped, droplet-shaped or globular viruses (Baquero, Y. Liu, et al., 2020), and bacterial viruses of the family *Plasmaviridae* (Krupovic and ICTV Report C, 2018), could represent additional realms. A variety of protein-coding cc-cRNAs have been discovered in metagenomes, and some of these could be independently originated viruses that would qualify as separate realms (B. D.

Lee et al., 2023). However, the available information on the virion organization, diversity and evolution of all these (putative) viruses remains scarce and thus their higher-level taxonomic assignment appears premature.

Can we expect to discover viruses representing additional new realms? It is likely that all major, widespread virus realms have been already sampled, a monumental achievement from a century of culture-dependent and more recent, culture-independent virus discovery efforts. However, it is hard to put an upper bound on the number of independently evolving smaller virus groups restricted to particular hosts or environments. We predict that sampling of the less explored host groups, such as archaea or protists, as well as extreme environments, will continue to expose previously unseen viruses, some of which will represent new realms. Classical, culture-dependent virus isolation will play a key role in this endeavor, whereas metagenomics and metatranscriptomics will be instrumental both in the preceding discovery phase, to identify candidates for new realms, and in the subsequent stage, for determining the extent of diversity and distribution of these new virus groups.

5. The key trends and processes of virus evolution

5.1. *Pronounced modularity of viruses*

Phylogenomic analysis of the rapidly growing collection of viral genomes allows us to infer some general trends and patterns of virus evolution (Koonin, Dolja and Krupovic, 2022). Any viral genome, with the exception of some satellite viruses and some derived viruses that have lost structural proteins, combines at least two functional modules, those encoding proteins involved in virion morphogenesis and those involved in genome replication and expression. Each of these modules can be represented by only a single gene (or even a ribozyme as in ribozymvirians) or can be highly complex, including numerous genes. For instance, many viruses of the realm *Floreoviria*, with small ssDNA genomes, e.g., circovirids, encode one capsid protein and one replication-associated protein (Rep) (Varsani et al., 2024), whereas large dsDNA viruses of both *Varidnaviria* and *Duplodnaviria* often encode a (nearly) complete DNA replisome

(Kazlauskas, Krupovic and Venclovas, 2016), a variety of proteins involved in transcription and translation (Grimm et al., 2019; Hillen et al., 2019; R. A. L. Rodrigues, da Silva, et al., 2020), and scores of structural proteins and virion assembly factors (Shao et al., 2022).

Nearly two decades ago, Dennis Bamford proposed the concept of “viral self”, that is, an evolutionary stable core of vertically inherited genes that define viral lineages (Krupovic and Bamford, 2008; Bamford, 2003). In the context of the assemblage of viruses now known as *Varidnaviria*, the self was defined as the DJR-MCP and the packaging ATPase. The concept has stood the test of time: today, we can indeed trace vertical evolution of viral realms through the phylogenies of VHGs that comprise the viral self. Notably, however, the types of VHGs comprising the self differ across the realms. In viruses of the realms *Riboviria*, with their small genomes, the self consists of the VHGs involved in replication, namely, RdRP and RT. By contrast, in the realms of viruses with DNA genomes, it is the morphogenesis module that qualifies as the self. There is a simple functional and evolutionary logic to this distinction. Viruses with RNA genomes cannot fully rely on the host replication machinery because RNA replication is not a normal cellular process. Therefore, these viruses must employ dedicated replication enzymes that cannot be replaced by cellular counterparts and thus are consistently inherited throughout the evolution of the viral realm (as always in biology, there are exceptions: ribozvirians and possibly other viroid-like viruses do not encode any replication proteins, instead hijacking the host transcription apparatus; these unusual viruses, however, encompass ribozymes required for replication). Conversely, RNA viruses often exchange and occasionally reinvent genes for structural proteins (e.g., capsid protein of togavirids or nucleocapsid protein of ortervirals (Koonin, Dolja and Krupovic, 2022)) so that these genes do not qualify as the viral self. In principle, the same logic could be applied to ssDNA viruses because cellular DNA replication mechanisms do not produce ssDNA intermediates and thus the virus must provide its own solution. However, unlike in the case of RNA and RT viruses, there are multiple unrelated enzymes that can yield ssDNA intermediates during replication, including several families of HUH endonucleases, Rep_trans endonucleases,

protein-primed family B DNA polymerases and a few others, and all of these are interchangeably used by ssDNA viruses from different realms (Figure 3B), as discussed above. Furthermore, in viruses of the realm *Pleomoviria*, there is certain flexibility with regard to which replicative intermediate, ssDNA or dsDNA, is packaged into the virion (Roine et al., 2010), which licenses exploration of a broader range of replicative strategies. Of the four realms formerly included in the *Monodnaviria*, only in *Floreoviria*, the replication module comprises the self and is used for tracing the evolutionary relationships between the member viruses although some of these viruses (e.g., anelovirids and bidnavirids) lack the signature two-domain HUH endonuclease (Kazlauskas, Varsani, et al., 2019; Krupovic, Varsani, et al., 2020). By contrast, viruses with dsDNA genomes can use the cellular replication machinery, and accordingly, these viruses widely differ in their repertoires of replication genes, from retaining the entire replication machinery to shedding it all. Furthermore, replication genes of these viruses, for example, DNA polymerases, are occasionally exchanged with cellular or non-orthologous or even non-homologous viral counterparts (Yutin, Tolstoy, et al., 2024). Therefore, the self of the respective realms can be only identified with genes encoding structural and morphogenetic proteins. A transitional state between viruses that require a specialized replication machinery and those that can exploit the host replication apparatus is represented by polyomavirids and papillomavirids. After switching from ssDNA to dsDNA genomes, these viruses lost the ancestral capacity to replicate via RCR due to mutational inactivation of the HUH endonuclease and switched to the theta replication mechanism typical of host DNA (Figure 3B), while retaining the inactivated endonuclease as a DNA-binding domain involved in the replisome function (Enemark et al., 2000; D. Li et al., 2003; Bochkareva et al., 2006).

5.2. *Complexification and functional convergence: major trends in viral evolution*

Reconstruction of virus genome evolution by mapping gene repertoires on phylogenetic trees of VHGs revealed the predominant trend of complexification by gene accretion although reductive evolution also appears to have occurred in some lineages (Koonin and Yutin, 2019; Wolf, Kazlauskas, et al., 2018; Koonin

and Yutin, 2018). Remarkably, this evolutionary trend is typical of both viruses with small genomes, such as ribovirians, and those with the largest known genomes, such as members of phylum *Nucleocytoviricota* in realm *Varidnaviria*, as well as caudoviricetes. Many virus lineages, however, appear to maintain relative stasis with respect to genome size and complexity across long evolutionary spans, exchanging and replacing many genes while maintaining nearly constant genome size. This evolutionary regime is clearly observed, for example, in phylum *Preplasmiviricota* within *Varidnaviria*, and could be linked to the physical constraints, namely, internal volume and packing capacity of the virions.

The great majority of the genes acquired by viruses come from the hosts, although notable cases of viral gene duplication followed by diversification are known as well (Senkevich et al., 2021), whereas some viral genes might have evolved de novo from non-coding sequences (Legendre, Fabre, et al., 2018; Legendre, Alempic, et al., 2019). The genes acquired by viruses encode four broad functional categories of proteins: (i) components of replication and expression systems enabling the relative autonomy of viruses from the respective cellular processes, (ii) structural proteins that initially become minor components of the virions but eventually can replace hallmark MCPs, for example, in pandoraviruses (Krupovic, Yutin, et al., 2016), (iii) metabolic enzymes that supplement and/or modulate the respective metabolic pathways of the host cells, (iv) proteins involved in various types of virus–host interactions, in particular, inhibitors of host immune systems. Because the benefits conferred by these functions are common to many if not all viruses, there is extensive convergence among the suites of genes captured by viruses of distant lineages, including different realms, especially in the first of the above functional categories. A notable case in point is the independent acquisition of helicases of three different superfamilies by orthornaviraens of different phyla (Wolf, Kazlauskas, et al., 2018). Apparently, helicases enable replication of (relatively) large RNA genomes, as indicated by the fact that viruses with positive-sense RNA genomes larger than 6 kb encode a helicase of one or another superfamily (Gorbalenya and Koonin, 1989), and thus open up the path of evolution toward genome complexification (again, there is a notable exception, a flavi-like ribovirus with a

nearly 12 kb genome lacking any helicase (Dahan et al., 2022)). A further increase of the RNA genome size, beyond the ~25 kb threshold, apparently required the acquisition of an enzymatic complex involved in RNA proofreading, which occurred in coronavirids and other nidovirals, the giants of the RNA domain of the virosphere (Neuman et al., 2025; Zhong et al., 2025; Gorbalenya, Enjuanes, et al., 2006). Similarly, many lineages of orthornaviraens independently acquired proteases of different families involved in virus polyprotein processing, yielding mature viral proteins which is one of the main strategies for expression in viruses of eukaryotes where initiation of translation typically occurs at the 5′-end of an mRNA (Koonin, Dolja and Krupovic, 2022; Koonin and Dolja, 1993).

A case that has been widely discussed in scientific and even popular literature, giving rise to provocative ideas, is the presence of genes encoding multiple components of the translation system, both proteins and tRNAs, in some viruses of phylum *Nucleocytoviricota*, in particular, mimivirids (R. A. L. Rodrigues, da Silva, et al., 2020). One of the mimivirids, tupanvirus, encodes the entire suite of translation system components, except for those of the ribosome (Abraham et al., 2018; R. A. L. Rodrigues, T. S. Arantes, et al., 2019). Given that translation is often considered a quintessential cellular functionality of which viruses are incapable, these findings triggered ideas on the evolution of the “giant” viruses from cells, possibly, even from a hypothetical fourth domain of life (Claverie, 2006; Raoult, Audic, et al., 2004; Colson, Gimenez, et al., 2011; Colson, de Lamballerie, et al., 2012; Colson, Levasseur, et al., 2018; Nasir et al., 2012; Nasir et al., 2017) (see also discussion below). Detailed phylogenetic analyses, however, unequivocally indicate that translation system components have been independently acquired by nucleocytoviricetes from eukaryotic hosts at different stages of evolution (Moreira and Brochier-Armanet, 2008; Williams et al., 2011; Yutin, Wolf, et al., 2014; Moreira and Lopez-Garcia, 2015). Indeed, some of these genes encoding orthologous proteins with the same function apparently have been captured independently by distantly related viruses (Yutin, Wolf, et al., 2014). Scattered genes for translation system components are present also in some bacterial and archaeal caudoviricetes, including multiple tRNAs (up to 62), tRNA synthetases and even

ribosomal proteins, clearly, convergent acquisitions (Y. Liu, Demina, et al., 2021; Al-Shayeb, Sachdeva, et al., 2020; Sencilo et al., 2013; Mizuno et al., 2019; L. X. Chen et al., 2022). The roles of virus-encoded translation system components in virus replication are not well understood, but it appears likely that viruses reprogram the host translation system to different degrees depending on the virus–host combinations, including the localization of the translation of viral mRNAs to a distinct compartment within the infected cell (R. Zhang et al., 2026). In addition, the role of certain viral tRNAs in countering diverse tRNA-targeting defense systems has also been demonstrated (van den Berg and Brouns, 2025).

In the second functional category, metabolic enzymes acquired by viruses, a notable case of convergence is the independent capture by viruses with large dsDNA genomes within both *Varidnaviria* and *Duplodnaviria* of enzymes and entire metabolic pathways for the biosynthesis of nucleotides, for example, thymidine kinase, thymidylate synthases of two unrelated families, and ribonucleotide reductases (Koonin, Dolja and Krupovic, 2022). Other metabolic pathways are limited to specific viral lineages, for example, the glycosphingolipid biosynthesis pathway in *Coccolithoviridae*, a family within *Nucleocytoviricota* (Vardi et al., 2009; Nissimov et al., 2019).

Which genes are retained by viruses upon accidental acquisition, presumably, depends on the host biology and response to virus infection, particularly in the case of metabolic enzymes. In some well-studied cases, the selection factors driving the fixation of the acquired genes in viruses are clear as, for example, for phages infecting cyanobacteria (Puxty et al., 2015). Many of these cyanophages encode the entire photosystem I or II, and some even both photosystems (Sharon et al., 2009; Fridman et al., 2017). Upon infection, the viral photosystems supplement the host ones, augmenting their activity so that the host cells remain energy-rich and conducive to massive replication of the cyanophages. In most cases, however, the connection with the host biology remains elusive, remaining to be elucidated through direct experiments with the respective virus–host systems, which are typically technically challenging if feasible at all.

The viral genes directly involved in virus–host interactions are even more specifically linked to the

host biology. A striking example is the movement proteins (MP) encoded by a broad variety of plant viruses including diverse orthornaviraens, pararnaviraens and cressdnaviricots (realm *Floreoviria*), and enabling intercellular movement of viruses in plants through the plasmodesmata (Mushegian and Elena, 2015). The MP originally evolved via a duplication of the SJR-MCP in an orthornaviraen, and then, spread extremely widely among plant viruses via HGT (Butkovic, Dolja, et al., 2023). Analogously, in animals, one of the principal mechanisms of virus entry into the host cells is membrane fusion, and accordingly, fusion proteins have spread across a broad variety of animal RNA and DNA viruses (Kielian and Rey, 2006; Guardado-Calvo and Rey, 2021).

The great majority (most likely, all) viruses with large genomes, that is, those of realms *Varidnaviria*, *Duplodnaviria*, and *Adnaviria*, and even some viruses with small genomes in *Riboviria* and some groups of ssDNA viruses encode proteins and often entire systems involved in the inhibition of host immunity (Crespo-Bellido and Duffy, 2023; S. Zhang et al., 2025). In large viruses, these genes can comprise a major fraction of the gene repertoire, and this part of the genome is highly variable, rarely being conserved beyond a viral family. The counter-defense genes can be conceptually divided into two large classes: (i) repurposed components of immune systems that function as dominant-negative inhibitors according to the “guns for hire principle” (Koonin, Makarova, Wolf and Krupovic, 2020), that is, the use of homologous components for both defense and counter-defense, and (ii) dedicated inhibitors, often without detectable homologs, some of which also employ the dominant-negative mechanism. The typical case in the first category are poxviruses, large animal viruses within *Nucleocytoviricota* which encode a broad variety of homologs of host proteins involved in immunity and related signal transduction pathways. Examples include homologs of tumor necrosis factor, interleukin 18, MHC Class I and apoptotic factor Bcl2, all of these represented by families of paralogs in poxvirus genomes (Senkevich et al., 2021). Although not all the mechanisms have been studied in detail, these proteins appear to act as dominant negative inhibitors of the cognate host immune pathways. In addition, poxviruses encode a family of proteins of the second class of counter-defense factors containing a chemokine-binding

domain that appears to have no cellular counterpart and thus seems to have evolved within poxviruses themselves, possibly, via a radical rearrangement of a pre-existing ancestral domain (C. A. Nelson et al., 2015). In other viruses of eukaryotes, both nucleocytoviricots, and mirusviricots and herpesvirals, the counter-defense genes are not well characterized, in particular, because few if any homologs of host immune proteins that could act as decoys have been identified. However, numerous genes of these viruses (with the partial exception of othoherpesvirids) remain orphans without detectable homologs, and many of these are likely to be involved in counter-defense. The difficulty with the identification of those is, in large part, that the hosts of many of these viruses remain unknown, and even for known unicellular eukaryotic hosts, the immune systems remain largely uncharacterized.

The counter-defense landscape of caudoviricetes as well as adnavirians is better known. It appears that these viruses employ primarily distinct counter-defense proteins, most without detectable homologs outside narrow groups of viruses, rather than mimics of host immune proteins. Most of these counter-defense proteins, in particular, anti-CRISPR proteins that have been investigated in detail in the last few years as part of the CRISPR boom (Pawluk et al., 2018; Bondy-Denomy, Maxwell, et al., 2023) are quite small (about or less than 100 amino acids) and often are not predicted to adopt a globular conformation or predicted to adopt a unique fold (Sahakyan et al., 2023), suggesting the possibility of their emergence *de novo* from non-coding sequences. Most of these small counter-defense proteins are highly specific towards particular bacterial or archaeal defense systems, binding to unique sites of immune proteins, such as, for example, CRISPR effector nucleases, including Cas9 and others. A notable exception to this high specificity are DNA-mimicking counter-defense proteins, small, negatively charged proteins that can inhibit host immune systems promiscuously by blocking their DNA-binding domains, that is, employing a distinct variant of the decoy strategy (H. C. Wang et al., 2019; S. Li et al., 2024). Although unique counter-defense proteins prevail in large viruses of prokaryotes, the guns for hire principle is employed as well. Thus, some phages, in particular jumbo phages with their large genomes, have coopted complete, functional CRISPR systems

targeting host defense systems or other MGE (Seed et al., 2013; Al-Shayeb, Skopintsev, et al., 2022). Even more common in the genomes of phages and archaeal viruses are CRISPR microarrays exploiting the host Cas proteins primarily to target other viruses (Faure et al., 2019; Medvedeva, Y. Liu, et al., 2019) and small anti-CRISPR RNAs that are copies of single CRISPR repeat and form non-productive complexes with Cas proteins, acting therefore as CRISPR-RNA decoys (Faure et al., 2019; Camara-Wilpert et al., 2023; Hayes et al., 2025).

Among viruses with small genomes, only some, apparently, a minority encode dedicated counter-defense proteins, again in connection with the host biology. Thus, diverse plant viruses as well as certain animal viruses, some with very small genomes, encode silencing suppressors, small proteins, possibly, emerging *de novo*, that inhibit the powerful RNA interference systems of the hosts (Baulcombe, 2022; W. X. Li and Ding, 2022).

To summarize this discussion of the general trends in the evolution of the virosphere, all viruses have a small core (self) that consists of a few, and in many cases, only one VHG(s) essential for viral replication and/or morphogenesis. These viral selves evolve vertically through eons, for millions or even billions of years, providing for the identification of viral realms, kingdoms and phyla. The rest of the viral genomes is highly dynamic and malleable, evolving through numerous gene gains, losses and exchanges, the exact history of which is often hard to reconstruct. There is, however, a lot of convergence and parallelism in the evolution of genes contributing to viral genome replication and expression due to shared requirements of diverse viruses. There is also some but lesser convergence among viral metabolic genes, whereas genes encoding virion components as well as those involved in counter-defense are an epitome of diversity and dynamism.

A consequence of the convergence in the acquisition of functionally analogous and often homologous host genes by diverse viruses is that, although viral realms, by definition, represent assemblages of viruses of independent origins, different realms are loosely connected by patchy sharing of homologous proteins. Prominent examples include the aforementioned RNA and DNA helicases (in particular, superfamily 3 helicases), DNA polymerases and DNA-dependent RNA polymerases, chymotrypsin-like and

papain-like proteases, movement proteins of plant viruses, and more (Koonin, Dolja, Krupovic, Varsani, et al., 2020).

6. Viruses as symbionts of cellular life forms

The pre-eminent immunologist Sir Peter Medawar famously said in his 1960 Nobel lecture that “No virus is known to do good. It has been well said that a virus is a piece of bad news wrapped up in protein” (Medawar, 1960). However, the more deeply is the virosphere explored, the clearer it becomes that the news conveyed by viral genomes is not necessarily bad, but perhaps, more often, is neutral or even good. Put another way, viruses are ubiquitous symbionts of cellular life forms, and their relationships with the hosts span the entire symbiotic spectrum, from parasitism to commensalism to mutualism. For obvious reasons, parasitic viruses that lyse host cells (thus, by definition, killing a unicellular host, and often, a multicellular one as well) have been discovered and studied first. However, already the early discovery of temperate (lysogenic) phages that can be vertically transmitted across many generations of bacteria as prophages showed that viruses are not invariably deleterious to the host (Lwoff, 1953). Indeed, prophages often encode defense systems and can have beneficial effects on the host fitness by protecting bacteria from infection by other, virulent phages (superinfection exclusion) (Bondy-Denomy, Qian, et al., 2016; Kirchberger, Martinez, Luker, et al., 2021; Brenes and Laub, 2025). There can be other benefits as well, notably transduction, that is, HGT for which phages serve as vehicles (E. Harrison and Brockhurst, 2017). The genes transferred by phages include defense systems, as per the “guns for hire” concept, antibiotic resistance genes as well as genes encoding enzymes of various metabolic pathways (Chevallereau and Westra, 2022; Y. Zhang et al., 2022; Fillol-Salom et al., 2022; Y. Zhao et al., 2025). Moreover, some bacteria and archaea have domesticated phages, turning them into Gene Transfer Agents (GTA), which are defective phages that incorporate into virions random host genes rather than the phage genome and function as dedicated, stress-induced HGT vehicles (Lang et al., 2012; Kogay and Zhaxybayeva, 2024). Notably, GTAs can rescue recipient cells from DNA damage by providing templates for homologous recombination (Gozzi

et al., 2022). Conceptually similar exaptation of large DNA viruses for the transfer of immunosuppression genes took place in parasitoid wasps, apparently on several independent occasions, leading to the emergence of viriforms, virus derivatives carrying host DNA (Drezen et al., 2022; Petersen et al., 2022; Kuhn and Koonin, 2023). Recently, it has been shown that a variety of phages considered virulent can accumulate in bacterial cells without either lysing them or integrating (Perfilyev et al., 2026; Dougherty et al., 2026), suggesting that viral commensalism is far more widespread than previously suspected.

A notable case of different types of symbiosis is presented by virophages, the viral symbionts of nucleocytoviricots (Koonin, Fischer, et al., 2024). Some of the virophages efficiently inhibit the propagation of the cognate large viruses, in effect functioning as adaptive immunity systems—and hence mutualists—for the unicellular eukaryotic host (Fischer and Hackl, 2016; Koslova et al., 2024). Other virophages have little if any effect on the supporting large virus and appear to be commensals. More generally, intervirus symbiosis and coevolution in complex virus–host systems are an important, currently poorly understood aspect of the evolution of the virosphere (Del Arco et al., 2024).

Every organism apparently supports a “healthy” virome, that is, commensal and mutualist viruses (Koonin, Dolja and Krupovic, 2021). For example, exploration of plant and fungal transcriptomes reveals an increasing number of cryptic viruses that cause no symptoms and have been largely missed by traditional virus identification approaches (Roossinck and Bazan, 2017; K. B. Gilbert et al., 2019). Similarly, animals, including humans, harbor a variety of viral commensals, for example, anellovirids with tiny ssDNA genomes that are ubiquitous and abundant in humans but have not been associated with any pathology (Butkovic, Kraberger, et al., 2023; Kaczorowska and van der Hoek, 2020; Kraberger et al., 2026). Both mammals and insects also host true viral mutualists, such as endogenous retroviruses (often denoted LTR retrotransposons) integrated in the genome, some of which form particles transferring RNA between host cells and apparently contributing to human physiology, in particular, that of the nervous system (Ashley et al., 2018; Pastuzyn et al., 2018; Shepherd, 2018; Segel et al., 2021).

In general, the scale of viral commensalism and mutualism in the biosphere remains to be elucidated. Infected cell lysis and host killing are not per se targets of selection in the evolution of viruses—rather, the maximum level of virus propagation is. Some viruses achieve this by killing the host, but for many, perhaps the majority of viruses, this is not the case.

7. The origins of viruses: genetic parasites as an intrinsic feature of life, the primordial replicator pool and exaptation of cellular genes

As emphasized above, viruses are literally ubiquitous in the biosphere. Apparently, all organisms, with the possible exception of some intracellular symbiotic bacteria with highly reduced genomes, are infected by multiple viruses, often highly diverse ones, representing different realms. Humans, for example, are hosts to numerous viruses of both kingdoms of *Riboviria*, *Floreoviria* (parvovirids, anellovirids, papillomavirids and others), *Varidnaviria* (poxvirids) and *Duplodnaviria* (orthohespervirids). The empirical data on the ubiquity of viruses are complemented by a theoretical argument on the inevitability of the emergence of genetic parasites in any replicator system (Koonin, Wolf, et al., 2017). This argument has been developed in thermodynamic terms, based on the entropy increase associated with the emergence of genetic parasites. In qualitative terms coming from game theory, the inevitability of the emergence of genetic parasites can be explained quite simply: as soon as, in a replicator system, there is an alienable resource that can be appropriated, such as a replicase, cheaters will emerge that will use that resource without making it. These theoretical considerations are, in turn, compatible with the reconstruction of the virome of the LUCA that is estimated to have existed about 4 billion years ago (Moody et al., 2024) and is the earliest point in the history of life for which such a reconstruction is attainable. This reconstruction indicates that the virome of the LUCA was already highly complex, comparable with the extant bacterial viromes, that is, probably included representatives of all the major viral realms (Krupovic, Dolja, et al., 2020).

The LUCA was certainly not the first life form to appear on earth but rather a product of extensive

evolution from primordial protocells that likely existed within the framework of the RNA world, where RNA molecules performed both template and catalytic functions (W. Gilbert, 1986; Joyce, 2002). The inevitability of genetic parasites implies that they emerged concomitantly with the very first replicators. Furthermore, whereas the primordial replicator pool is thought to initially have consisted of RNA molecules only, small and then larger DNA replicators must have evolved at subsequent stages of evolution, long before the time of the LUCA. Thus, replicators corresponding to all Baltimore classes of extant viruses were most likely already parts of the primordial pool (Koonin, Senkevich, et al., 2006). Comparative analysis of the essential proteins involved in viral genome replication is compatible with the very early origin of the viral replication machineries (Krupovic, Dolja, et al., 2019). Indeed, behind the diversity of viral replication enzymes, there is remarkable, deep unity. The replication enzymes of most viruses from different realms, including RdRP, RT, HUH endonuclease, archaeo-eukaryotic primases and family B and A DNAPs, all contain homologous core RNA Recognition Motif (RRM) fold domains. The RRM is one of the most common nucleic acid binding domains in nature, found in an enormous variety of proteins, most of them devoid of enzymatic activity (Clery et al., 2008). The parsimonious explanation of the presence of catalytically active RRM in viral replication enzymes is that they all radiated from a non-enzymatic ancestral RRM that was likely one of the earliest protein domains to evolve and could have served as a cofactor to ribozyme polymerases at the exit from the RNA world stage (Krupovic, Dolja, et al., 2019) (Figure 8). The second widespread structural scaffold found in diverse RNA and DNA polymerases is the double psi-beta barrel domain that has been suggested to represent the ancestral replicase of cellular organisms (Sauguet, 2019; Koonin, Krupovic, Ishino, et al., 2020).

Thus, all types of replicators, including parasitic ones, in all likelihood, evolved very early in the evolution of life, prior to the advent of modern-type cells (Figure 8). However, parasitic replicators are not (yet) viruses, and bona fide viruses hardly could have existed before full-fledged cells emerged. Indeed, for major virion proteins, cellular ancestors have been identified, the most prominent cases in point being the origin of SJR-MCP and DJR-MCP from distinct

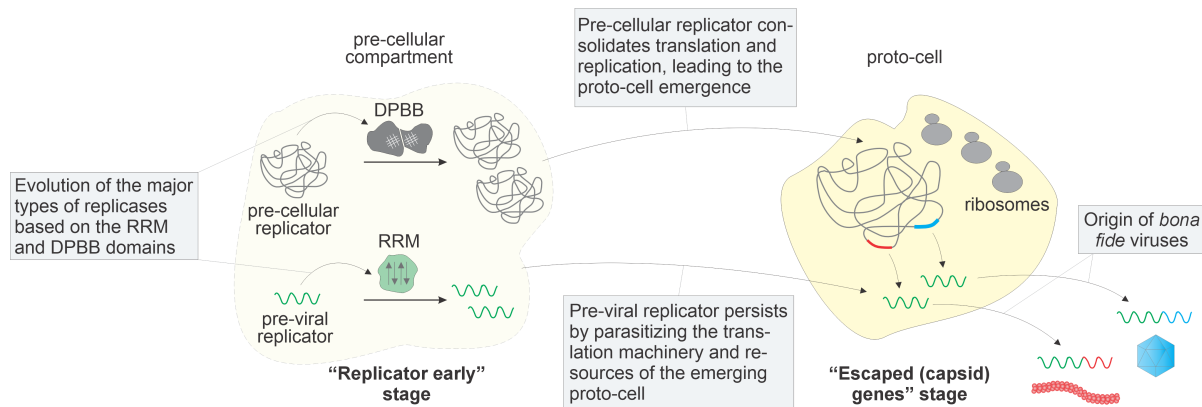


Figure 8. Origin of viruses. The depicted scenario combines origin of viral genomes and replication machinery from the primordial (pre-cellular) pool of replicators with subsequent exaptation of cellular proteins, such as carbohydrate-binding jelly-roll proteins, as capsid proteins, resulting in the emergence of bona fide viruses. Replicative genes are rendered in green, and structural genes, in red and blue. The two major types of primordial replicases based on the RNA recognition motif (RRM) and double psi-beta barrel (DPBB) domain are hypothesized to have evolved in pre-viral and pre-cellular replicators, respectively.

families of sugar-binding proteins and glycoside hydrolases, respectively (Krupovic and Koonin, 2017b; Krupovic, Makarova, et al., 2022), suggesting viruses appeared on the scene at a relatively advanced stage of evolution, after multiple protein families have already diversified, even if long before the LUCA (Figure 8).

For decades, the origin of viruses has been discussed in terms of three distinct scenarios: (i) viruses early, preceding the origin of cells, (ii) viruses late, via reductive evolution of cellular parasites, (iii) escaped genes, that is, another version of viruses late, deriving viruses from “regular” cellular genes attaining replicative autonomy (Koonin, Senkevich, et al., 2006; Forterre and Prangishvili, 2009; Luria and Darnell, 1967; Forterre, 2006). The evolutionary scenario outlined above is a hybrid of “viruses early” and “escaped genes”: here, the viral replication machinery derives from the primordial pool of replicators (viruses early) whereas the virions evolve later from cellular proteins (escaped genes) (Figure 8). Importantly, the origin of viruses from non-viral symbiotic replicators, in particular, plasmids, is not limited to the hypothetical scenario of viral origin. Rather, this path of evolution was recapitulated on multiple occasions, in particular, at the origins of several groups of viruses with ssDNA genomes (Kazlauskas, Varsani, et al., 2019) as well as the origin of the expansive orthornaviraen family *Botourmiaviridae* from capsidless narnavirids (Wolf, Kazlauskas, et al., 2018).

8. Origins of the eukaryotic virome

The origin of eukaryotes (eukaryogenesis) is the second pivotal major transition in the evolution of life, after the origin of cells (Maynard Smith and Szathmary, 1995; Szathmary, 2015). According to the latest, best supported scenario, the eukaryotic cell evolved as a result of engulfment by an Asgard archaeon related to the order Hodarchaeales of an alphaproteobacterium that became the protomitochondrial endosymbiont (Eme et al., 2023; Tobiasson et al., 2026) (Figure 9A). The reconstruction of the virome of the Last Eukaryotic Common Ancestor (LECA) showed, not unexpectedly, that this ancestral virome was highly complex, including representatives of all the larger realms, and likely, most of the phyla of the known viruses of eukaryotes (Krupovic, Dolja, et al., 2023). The unexpected outcome of this reconstruction, however, was the apparent origin of all major groups of viruses of eukaryotes from bacterial rather than archaeal ancestors (Figure 9A,B). In particular, all viruses of Asgard archaea identified to date are typical archaeal viruses without traceable links to viruses of eukaryotes (Tamarit et al., 2022; Medvedeva, Sun, et al., 2022; Rambo et al., 2022). Given that viruses are genetic parasites and the eukaryotic systems for information storage and transmission are of Asgard archaeal descent almost in their entirety (Tobiasson et al., 2026), this finding appeared puzzling and even paradoxical. The best

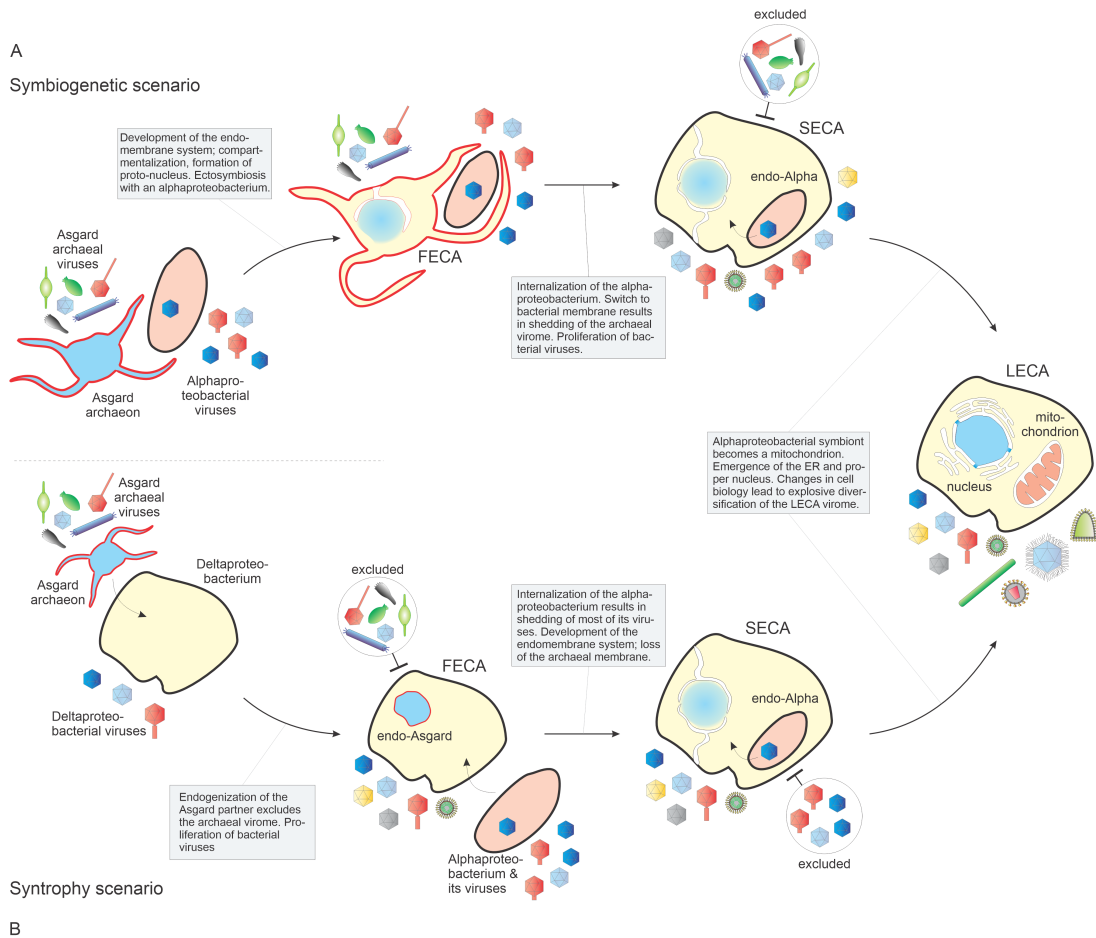


Figure 9. Eukaryogenesis and the origin of the eukaryotic virome. Two alternative scenarios of eukaryogenesis are shown. (A) Symbiogenetic scenario whereby an Asgard archaeon with a complex cellular organization engulfed an alphaproteobacterium, the proto-mitochondrial endosymbiont, with subsequent replacement of the archaeal membrane (red outline) by the bacterial one (black outline). (B) Syntrophy scenario including two endosymbioses whereby a bacterium (possibly, a deltaproteobacterium) engulfed an Asgard archaeon, resulting in the dissolution of the archaeal membrane and followed by a secondary engulfment of an alphaproteobacterium. In this scenario, the continuity of the bacterial membrane is preserved throughout. Under both scenarios, archaeal viruses are excluded from the emerging eukaryote, and as a result, the eukaryotic virome is seeded by bacterial viruses. FECA, SECA and LECA: first, second and last eukaryotic common ancestor, respectively.

possible explanation seems to be the major difference between the structures of archaeal and bacterial membranes (Lombard et al., 2012), and likely membrane-embedded viral receptors that could have led to the exclusion of archaeal viruses from the evolving eukaryotic cells upon the replacement of the archaeal membranes (and cell walls) with bacterial ones (Krupovic, Dolja, et al., 2023). Under more complex scenarios of eukaryogenesis—such as the syntrophy model, which postulates the initial engulfment of an archaeon by a bacterium, preserving the continuity of bacterial membranes throughout

eukaryogenesis (Moreira and Lopez-Garcia, 1998; Lopez-Garcia and Moreira, 2020)—exclusion of archaeal viruses could have occurred at the initial symbiosis stage (Figure 9B).

9. Reproducers and replicators, fundamental virus-cell divide and the place of the virosphere within the replicator space

Evolving biological entities belong to two fundamentally distinct types: reproducers and replicators (Szathmari and Maynard Smith, 1997). Reproducers are,

essentially, analog devices that retain physical continuity throughout the course of evolution. All cellular life forms possess the properties of reproducers as reflected in the famous formula of Rudolf Virchow: *Omnia cellula e cellula*. Replicators, in contrast, are digital devices, so that physical continuity is not necessary for their propagation and evolution; information contained in the nucleotide sequence of the genome is sufficient. Clearly, all MGEs, viruses in particular, are replicators (the propagation of many viruses requires not only the genomic nucleic acid but also some virion protein(s) or even macromolecular structures to enter the cell; however, such structures are not transmitted to the next generation). Genomes of cellular life forms in themselves are replicators as well, so that organisms actually represent a union of reproducers and replicators. The establishment of mutualistic symbiosis between primordial reproducers and replicators might have been a pivotal point in the origin of life (Babajanyan et al., 2023).

The split between reproducers and replicators appears to be the most fundamental divide in biology. There is no evidence that the barrier between these two types of propagating, evolving entities has ever been crossed in the more than 4 billion years of the evolution of life, and given this perpetual but not blending symbiosis, this barrier appears to be fundamentally impenetrable. Attempts to define viruses as biological entities distinct from cells on the basis of simple criteria such as size have been hopelessly compromised by the latest findings, in particular the discovery of giant viruses (Koonin, Dolja, Krupovic and Kuhn, 2021). The definition of viruses as capsid-encoding organisms, in contrast to cellular life forms, which are ribosome-encoding organisms (Raoult and Forterre, 2008), fares better (regardless of whether or not it is appropriate to call viruses organisms). However, the discovery of viruses that encode almost all components of the translation system apart from the ribosome or multiple ribosomal proteins (see above) puts even this definition into doubt. What if a virus is discovered that actually encodes its own ribosome? Will this eliminate or blur the distinction between viruses and cells, as it has been repeatedly suggested giant viruses do? In our strong opinion, no such blurring will occur because even a ribosome-encoding virus will remain a replicator, not a reproducer. Thus, the best definition of a virus may be “a replicator en-

coding at least one protein encasing the viral genome or an MGE demonstrably derived from such a replicator” (Koonin, Dolja, Krupovic and Kuhn, 2021).

The definition of viruses as a distinct type of replicators prompts the question on the position of the virosphere within the space of replicators (ibid.). Apart from typical viruses that meet the above definition and comprise the “orthovirosphere”, a broad variety of virus-like MGE that can be derived from viruses or ancestral to viruses inhabit the “perivirosphere” (Figure 10). The denizens of the perivirosphere include capsid-less derivatives of regular viruses (for example, narnaviruses, mitoviruses, umbraviruses, endornaviruses and others within *Riboviria*), various satellite nucleic acids, viroids and viroid-like cccRNA replicators, and other MGE, such as viriforms. The orthovirosphere and perivirosphere are embedded in the greater replicator space that encompasses “non-virus-like” replicators such as typical transposons, integrating conjugative elements, plasmids as well as genomes of cellular life forms (Figure 10). Crucially, the boundaries between the different domains of the replicator space are porous, and evolutionary interconversions between different types of elements abound. In sharp contrast, as emphasized above, the wall separating the replicator space from reproducers appears to be impenetrable.

10. Concluding remarks

Obviously, in this rather long review article, we could only superficially cover most aspects of the organization and evolution of the vast world of viruses. The diversity of known viruses is now growing more rapidly than ever, primarily, through metagenome and metatranscriptome mining, often revealing previously unsuspected features, for example, viruses with genomes in excess of 4 mb or diverse (putative) viruses with cccRNA genomes. Nevertheless, however astonishing this might be, it appears that we already know the global organization of the virosphere and the main structure of the megataxonomy of viruses. At the same time, it should be emphasized that this structure is an inherently moving target, as demonstrated, in particular, by the ongoing splitting of some of the viral realms. From the structuring of the virosphere and extensive phylogenomic analyses, general principles of virus evolution are emerging and new research frontiers are opening up. Some of

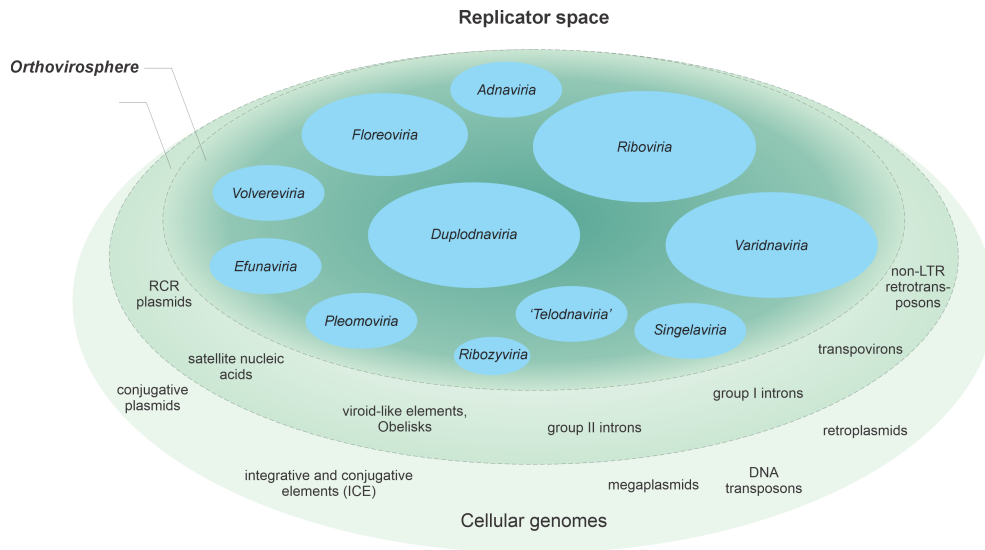


Figure 10. The replicators space and its three domains. The replicator space is represented as consisting of three domains: orthovirosphere including the 11 recognized realms of viruses; perivirosphere including virus-like replicators; and the outer replicator space including all other replicators, in particular, the genomes of cellular life forms. The boundaries between the domains are shown by dashed lines to emphasize multiple evolutionary interconversions between different types of replicators. Obelisks are cccRNA replicators encoding an uncharacterized protein and replicating in bacterial cells (Zheludev et al., 2024; Lopez-Simon et al., 2025; Urayama, Fukudome, Mutz, et al., 2026). Transpovirons are small dsDNA plasmids that are commensals of viruses in the phylum *Nucleocyotivicota* (Krupovic, Yutin, et al., 2016; Desnues et al., 2012).

the important directions of future research include the study of different facets of virus–host symbioses as well as coevolution of viruses in complex systems such as those including satellite viruses. There is no doubt that new, fascinating features of the virosphere and its evolution will be uncovered.

Acknowledgements

EVK is supported by the Intramural Research Program of the National Institutes of Health of the USA.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

References

- Abad, L., C. H. Gauthier, I. Florian, D. Jacobs-Sera and G. F. Hatfull, “The heterogenous and diverse population of prophages in *Mycobacterium* genomes”, *mSystems* **8** (2023), no. 5, article no. e0044623.
- Abeywickrama-Samarakoon, N., J. C. Cortay, C. Sureau, et al., “Hepatitis delta virus histone mimicry drives the recruitment of chromatin remodelers for viral RNA replication”, *Nat. Commun.* **11** (2020), no. 1, article no. 419.
- Abrahamo, J., L. Silva, L. S. Silva, et al., “Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere”, *Nat. Commun.* **9** (2018), no. 1, article no. 749.
- Abrescia, N. G., J. J. Cockburn, J. M. Grimes, et al., “Insights into assembly from structural analysis of bacteriophage PRD1”, *Nature* **432** (2004), no. 7013, pp. 68–74.
- Adams, M. J., E. J. Lefkowitz, A. M. King, et al., “50 years of the International Committee on Taxonomy of Viruses: progress and prospects”, *Arch. Virol.* **162** (2017), no. 5, pp. 1441–1446.
- Adriaenssens, E. M., R. Cook, V. V. Dolja, et al., “ICTV virus taxonomy profile: *Duplodnaviria* 2025”, *J. Gen. Virol.* **106** (2025), no. 9, article no. 002139.
- Agol, V. I., “Towards the system of viruses”, *Biosystems* **6** (1974), no. 2, pp. 113–132.
- Alarcon-Schumacher, T., D. Lucking and S. Erdmann, “Revisiting evolutionary trajectories and the organization of the Pleolipoviridae family”, *PLoS Genet.* **19** (2023), no. 10, article no. e1010998.
- Ashley, J., B. Cordy, D. Lucia, L. G. Fradkin, V. Budnik and T. Thomson, “Retrovirus-like Gag protein Arc1 binds RNA and traffics across synaptic boutons”, *Cell* **172** (2018), no. 1–2, pp. 262–274.
- Babajanyan, S. G., Y. I. Wolf, A. Khachatryan, A. Allahverdyan, P. Lopez-Garcia and E. V. Koonin, “Coevolution of reproducers and replicators at the origin of life and the conditions for

- the origin of genomes”, *Proc. Natl. Acad. Sci. USA* **120** (2023), no. 14, article no. e2301522120.
- Baltimore, D., “Expression of animal virus genomes”, *Bacteriol. Rev.* **35** (1971), no. 3, pp. 235–241.
- Bamford, D. H., “Do viruses form lineages across different domains of life?”, *Res. Microbiol.* **154** (2003), no. 4, pp. 231–236.
- Baquero, D. P., E. A. Bignon and M. Krupovic, “Pleomorphic viruses establish stable relationship with marine hyperthermophilic archaea”, *ISME J.* **18** (2024), no. 1, article no. wrae008.
- Baquero, D. P., Y. Liu, F. Wang, E. H. Egelman, D. Prangishvili and M. Krupovic, “Structure and assembly of archaeal viruses”, *Adv. Virus Res.* **108** (2020), pp. 127–164.
- Baquero, D. P., S. Medvedeva, C. Martin-Gallaussiaux, et al., “Stable coexistence between an archaeal virus and the dominant methanogen of the human gut”, *Nat. Commun.* **15** (2024), no. 1, article no. 7702.
- Bardy, P., C. I. W. MacDonald, P. C. Kirchberger, et al., “Penton blooming, a conserved mechanism of genome delivery used by disparate microviruses”, *mBio* **16** (2025), no. 4, article no. e0371324.
- Barreat, J. G. N. and A. Katzourakis, “Phylogenomics of the Maverick virus-like mobile genetic elements of vertebrates”, *Mol. Biol. Evol.* **38** (2021), no. 5, pp. 1731–1743.
- Bath, C., T. Cukalac, K. Porter and M. L. Dyal-Smith, “His1 and His2 are distantly related, spindle-shaped haloviruses belonging to the novel virus group, Salterprovirus”, *Virology* **350** (2006), no. 1, pp. 228–239.
- Baulcombe, D. C., “The role of viruses in identifying and analyzing RNA silencing”, *Annu. Rev. Virol.* **9** (2022), no. 1, pp. 353–373.
- Benning, F. M. C., S. Jenni, C. Y. Garcia, T. H. Nguyen, X. Zhang and L. H. Chao, “Helical reconstruction of VP39 reveals principles for baculovirus nucleocapsid assembly”, *Nat. Commun.* **15** (2024), no. 1, article no. 250.
- Bergner, L. M., R. J. Orton, A. Broos, et al., “Diversification of mammalian deltaviruses by host shifting”, *Proc. Natl. Acad. Sci. USA* **118** (2021), no. 3, article no. e2019907118.
- Bignon, E. A., K. R. Chou, E. Roine and N. D. Tischler, “Halorubrum pleomorphic virus-6 membrane fusion is triggered by an S-layer component of its haloarchaeal host”, *Viruses* **14** (2022), no. 2, article no. 254.
- Black, E. J., C. S. Powell, D. M. Dempsey, R. C. Hendrickson, L. R. Mims and E. J. Lefkowitz, “Virus taxonomy: the database of the International Committee on Taxonomy of Viruses”, *Nucleic Acids Res.* **54** (2026), no. D1, pp. D776–D789.
- Bochkareva, E., D. Martynowski, A. Seitova and A. Bochkarev, “Structure of the origin-binding domain of simian virus 40 large T antigen bound to DNA”, *EMBO J.* **25** (2006), no. 24, pp. 5961–5969.
- Boisvert, N., S. Thurmond, C. Elenberger, et al., “Anellovirus protein encoded by ORF2/3 functions as the viral replication initiation protein”, *Proc. Natl. Acad. Sci. USA* **122** (2025), no. 52, article no. e2516306122.
- Bondy-Denomy, J., K. L. Maxwell and A. R. Davidson, “Ten years of anti-CRISPR research”, *J. Mol. Biol.* **435** (2023), no. 7, article no. 168058.
- Bondy-Denomy, J., J. Qian, E. R. Westra, A. Buckling, D. S. Guttman, A. R. Davidson and K. L. Maxwell, “Prophages mediate defense against phage infection through diverse mechanisms”, *ISME J.* **10** (2016), no. 12, pp. 2854–2866.
- Brenes, L. R. and M. T. Laub, “E. coli prophages encode an arsenal of defense systems to protect against temperate phages”, *Cell Host Microb.* **33** (2025), no. 6, pp. 1004–1018.
- Buscaglia, M., M. Diallo, W. Davis, T. James, B. Diez and F. Schulz, “Megagenomoviruses with nearly five megabase genomes and ribosomal proteins”, preprint, 2025. Online at <https://www.researchsquare.com/article/rs-8000643/v1> (accessed on April 27, 2026).
- Butkovic, A., V. V. Dolja, E. V. Koonin and M. Krupovic, “Plant virus movement proteins originated from jelly-roll capsid proteins”, *PLoS Biol.* **21** (2023), no. 6, article no. e3002157.
- Butkovic, A., S. Kraberger, Z. Smeele, et al., “Evolution of anelloviruses from a circovirus-like ancestor through gradual augmentation of the jelly-roll capsid protein”, *Virus Evol.* **9** (2023), no. 1, article no. vead035.
- Cai, Z., C. Lu, J. He, et al., “Identification and characterization of circRNAs encoded by MERS-CoV, SARS-CoV-1 and SARS-CoV-2”, *Brief Bioinform.* **22** (2021), no. 2, pp. 1297–1308.
- Camara-Wilpert, S., D. Mayo-Munoz, J. Russel, R. D. Fagerlund, J. S. Madsen, P. C. Fineran, S. J. Sorensen and R. Pinilla-Redondo, “Bacteriophages suppress CRISPR-Cas immunity using RNA-based anti-CRISPRs”, *Nature* **623** (2023), no. 7987, pp. 601–607.
- Carding, S. R., N. Davis and L. Hoyles, “Review article: the human intestinal virome in health and disease”, *Aliment Pharmacol. Ther.* **46** (2017), no. 9, pp. 800–815.
- Caston, J. R. and J. L. Carrascosa, “The basic architecture of viruses”, *Subcell Biochem.* **68** (2013), pp. 53–75.
- Chamakura, K. R. and R. Young, “Single-gene lysis in the metagenomic era”, *Curr. Opin. Microbiol.* **56** (2020), pp. 109–117.
- Chandler, M., F. de la Cruz, F. Dyda, A. B. Hickman, G. Moncalian and B. Ton-Hoang, “Breaking and joining single-stranded DNA: the HUH endonuclease superfamily”, *Nat. Rev. Microbiol.* **11** (2013), no. 8, pp. 525–538.
- Chang, W. S., J. H. Pettersson, C. Le Lay, M. Shi, N. Lo, M. Wille, J. S. Eden and E. C. Holmes, “Novel hepatitis D-like agents in vertebrates and invertebrates”, *Virus Evol.* **5** (2019), no. 2, article no. vez021.
- Chen, L. X., A. L. Jaffe, A. L. Borges, P. I. Penev, T. C. Nelson, L. A. Warren and J. F. Banfield, “Phage-encoded ribosomal protein S21 expression is linked to late-stage phage replication”, *ISME Commun.* **2** (2022), no. 1, article no. 31.
- Chevallereau, A. and E. R. Westra, “Bacterial immunity: mobile genetic elements are hotspots for defence systems”, *Curr. Biol.* **32** (2022), no. 17, R923–R926.
- Chung, D., N. Brask, S. Matar, et al., “Persistent mirusvirus infection in the marine protist *Aurantiochytrium*”, *Nat. Commun.* **16** (2025), no. 1, article no. 9922.
- Ciccarelli, F. D., T. Doerks, C. von Mering, C. J. Creevey, B. Snel and P. Bork, “Toward automatic reconstruction of a highly resolved tree of life”, *Science* **311** (2006), no. 5765, pp. 1283–1287.
- Claverie, J. M., “Viruses take center stage in cellular evolution”, *Genome Biol.* **7** (2006), no. 6, article no. 110.
- Clery, A., M. Blatter and F. H. Allain, “RNA recognition motifs: boring? Not quite”, *Curr. Opin. Struct. Biol.* **18** (2008), no. 3, pp. 290–298.
- Colson, P., X. de Lamballerie, G. Fournous and D. Raoult, “Reclassification of giant viruses composing a fourth domain of life in the new order Megavirales”, *Intervirology* **55** (2012), no. 5, pp. 321–332.

- Colson, P., G. Gimenez, M. Boyer, G. Fournous and D. Raoult, “The giant Cafeteria roenbergensis virus that infects a widespread marine phagocytic protist is a new member of the fourth domain of Life”, *PLoS One* **6** (2011), no. 4, article no. e18935.
- Colson, P., A. Levasseur, B. La Scola, V. Sharma, A. Nasir, P. Pontarotti, G. Caetano-Anolles and D. Raoult, “Ancestrality and mosaicism of giant viruses supporting the definition of the fourth TRUC of microbes”, *Front. Microbiol.* **9** (2018), article no. 2668.
- Cornillez-Ty, C. T. and D. W. Lazinski, “Determination of the multimerization state of the hepatitis delta virus antigens in vivo”, *J. Virol.* **77** (2003), no. 19, pp. 10314–10326.
- Coulibaly, F., “Structure of the poxvirus core”, *Nat. Struct. Mol. Biol.* **31** (2024), no. 7, pp. 1001–1003.
- Creager, A. N. H., “Tobacco mosaic virus and the history of molecular biology”, *Annu. Rev. Virol.* **9** (2022), no. 1, pp. 39–55.
- Crespo-Bellido, A. and S. Duffy, “The how of counter-defense: viral evolution to combat host immunity”, *Curr. Opin. Microbiol.* **74** (2023), article no. 102320.
- Crick, F. H. and J. D. Watson, “Structure of small viruses”, *Nature* **177** (1956), no. 4506, pp. 473–475.
- Crump, C., “Virus assembly and egress of HSV”, *Adv. Exp. Med. Biol.* **1045** (2018), pp. 23–44.
- Dahan, J., Y. I. Wolf, G. E. Orellana, E. J. Wenninger, E. V. Koonin and A. V. Karasev, “A novel Flavi-like virus in Alfalfa (*Medicago sativa* L.) crops along the Snake River Valley”, *Viruses* **14** (2022), no. 6, article no. 1320.
- De Castro, E., C. Hulo, P. Masson, A. Auchincloss, A. Bridge and P. Le Mercier, “ViralZone 2024 provides higher-resolution images and advanced virus-specific resources”, *Nucleic Acids Res.* **52** (2024), no. D1, pp. D817–D821.
- De Colibus, L., E. Roine, T. S. Walter, et al., “Assembly of complex viruses exemplified by a halophilic euryarchaeal virus”, *Nat. Commun.* **10** (2019), no. 1, article no. 1456.
- De Koch, M. D., M. Krupovic, R. Fielding, et al., “Novel lineage of anelloviruses with large genomes identified in dolphins”, *J. Virol.* **99** (2025), no. 1, article no. e0137024.
- de la Higuera, I., G. W. Kasun, E. L. Torrance, et al., “Unveiling Crucivirus diversity by mining metagenomic data”, *mBio* **11** (2020), no. 5, article no. e01410-20.
- Del Arco, A., M. G. Fischer and L. Becks, “Evolution of exploitation and replication of giant viruses and virophages”, *Virus Evol.* **10** (2024), no. 1, article no. veae021.
- Demina, T. A., M. Dyal-Smith, M. Jalasvuori, S. Du and H. M. Oksanen, “ICTV virus taxonomy profile: Sphaerolipoviridae 2023”, *J. Gen. Virol.* **104** (2023), no. 3, article no. 001830.
- Desnues, C., B. La Scola, N. Yutin, et al., “Provirophages and transpovirons as the diverse mobilome of giant viruses”, *Proc. Natl. Acad. Sci. USA* **109** (2012), no. 44, pp. 18078–18083.
- DiMaio, F., X. Yu, E. Rensen, M. Krupovic, D. Prangishvili and E. H. Egelman, “Virology. A virus that infects a hyperthermophile encapsidates A-form DNA”, *Science* **348** (2015), no. 6237, pp. 914–917.
- Dion, M. B., F. Oechslin and S. Moineau, “Phage diversity, genomics and phylogeny”, *Nat. Rev. Microbiol.* **18** (2020), no. 3, pp. 125–138.
- Dokland, T., R. McKenna, L. L. Ilag, B. R. Bowman, N. L. Incardona, B. A. Fane and M. G. Rossmann, “Structure of a viral procapsid with molecular scaffolding”, *Nature* **389** (1997), no. 6648, pp. 308–313.
- Dolja, V. V., “How nidoviruses evolved the largest known RNA genomes”, *Proc. Natl. Acad. Sci. USA* **122** (2025), no. 11, article no. e2501153122.
- Doore, S. M. and B. A. Fane, “The microviridae: diversity, assembly, and experimental evolution”, *Virology* **491** (2016), pp. 45–55.
- Dotto-Maurel, A., I. Arzul, B. Morga and G. Chevignon, “Herpesviruses: overview of systematics, genomic complexity and life cycle”, *Virol. J.* **22** (2025), no. 1, article no. 155.
- Dougherty, P. E., C. Bernard, A. B. Carstens, et al., “Persistent virulent phages exist across bacterial isolates”, *Nat. Microbiol.* **11** (2026), no. 1, pp. 31–41.
- Drezen, J. M., A. Bezier, G. R. Burke and M. R. Strand, “Bacoviruses, ichtnoviruses, and virus-like particles from parasitoid wasps retain many features of their virus ancestors”, *Curr. Opin. Insect Sci.* **49** (2022), pp. 93–100.
- Edgar, R. C., J. Taylor, V. Lin, et al., “Petabase-scale sequence alignment catalyses viral discovery”, *Nature* **602** (2022), no. 7895, pp. 142–147.
- El Omari, K., S. Li, A. Kotecha, et al., “The structure of a prokaryotic viral envelope protein expands the landscape of membrane fusion proteins”, *Nat. Commun.* **10** (2019), no. 1, article no. 846.
- Eme, L., D. Tamarit, E. F. Caceres, et al., “Inference and reconstruction of the heimdallarchaeal ancestry of eukaryotes”, *Nature* **618** (2023), no. 7967, pp. 992–999.
- Enemark, E. J., G. Chen, D. E. Vaughn, A. Stenlund and L. Joshua-Tor, “Crystal structure of the DNA binding domain of the replication initiation protein E1 from papillomavirus”, *Mol. Cell* **6** (2000), no. 1, pp. 149–158.
- Faure, G., S. A. Shmakov, W. X. Yan, D. R. Cheng, D. A. Scott, J. E. Peters, K. S. Makarova and E. V. Koonin, “CRISPR-Cas in mobile genetic elements: counter-defence and beyond”, *Nat. Rev. Microbiol.* **17** (2019), no. 8, pp. 513–525.
- Fillol-Salom, A., J. T. Rostol, A. D. Ojiogu, J. Chen, G. Douce, S. Humphrey and J. R. Penades, “Bacteriophages benefit from mobilizing pathogenicity islands encoding immune systems against competitors”, *Cell* **185** (2022), no. 17, pp. 3248–3262.
- Fischer, M. G., “The virophage family Lavidaviridae”, *Curr. Issues Mol. Biol.* **40** (2021), pp. 1–24.
- Fischer, M. G. and T. Hackl, “Host genome integration and giant virus-induced reactivation of the virophage mavirus”, *Nature* **540** (2016), pp. 288–291.
- Forgia, M., B. Navarro, S. Daghino, et al., “Hybrids of RNA viruses and viroid-like elements replicate in fungi”, *Nat. Commun.* **14** (2023), no. 1, article no. 2591.
- Forterre, P., “The origin of viruses and their possible roles in major evolutionary transitions”, *Virus Res.* **117** (2006), no. 1, pp. 5–16.
- Forterre, P. and D. Prangishvili, “The origin of viruses”, *Res. Microbiol.* **160** (2009), no. 7, pp. 466–472.
- Forterre, P. and D. Prangishvili, “The major role of viruses in cellular evolution: facts and hypotheses”, *Curr. Opin. Virol.* **3** (2013), no. 5, pp. 558–565.
- Forterre, P., N. Soler, M. Krupovic, E. Marguet and H. W. Ackermann, “Fake virus particles generated by fluorescence microscopy”, *Trends Microbiol.* **21** (2013), no. 1, pp. 1–5.
- Fridman, S., J. Flores-Urbe, S. Larom, et al., “A myovirus encoding both photosystem I and II proteins enhances cyclic elec-

- tron flow in infected *Prochlorococcus* cells”, *Nat. Microbiol.* **2** (2017), no. 10, pp. 1350–1357.
- Gaia, M., L. Meng, E. Pelletier, et al., “Mirusviruses link herpesviruses to giant viruses”, *Nature* **616** (2023), no. 7958, pp. 783–789.
- Gan, T. and D. Wang, “Picobirnaviruses encode proteins that are functional bacterial lysins”, *Proc. Natl. Acad. Sci. USA* **120** (2023), no. 37, article no. e2309647120.
- Gebhard, L. J., Y. Tomaru, K. Okamoto and A. Munke, “Structural diversity and conservation among CRESS-DNA bacilladnaviruses revealed through cryo-EM and computational modelling”, *Viol. J.* **22** (2025), no. 1, article no. 391.
- Gilbert, K. B., E. E. Holcomb, R. L. Allscheid and J. C. Carrington, “Hiding in plain sight: new virus genomes discovered via a systematic analysis of fungal public transcriptomes”, *PLoS One* **14** (2019), no. 7, article no. e0219207.
- Gilbert, W., “The RNA world”, *Nature* **319** (1986), p. 618.
- Gladyshev, E. A. and I. R. Arkhipova, “A widespread class of reverse transcriptase-related cellular genes”, *Proc. Natl. Acad. Sci. USA* **108** (2011), no. 51, pp. 20311–20316.
- Gong, Z. and G. Z. Han, “Insect retroelements provide novel insights into the origin of hepatitis B viruses”, *Mol. Biol. Evol.* **35** (2018), no. 9, pp. 2254–2259.
- Gong, Z. and G. Z. Han, “The cryptic diversity of hepadnavirus relatives”, *mBio* **16** (2025), no. 12, article no. e0254125.
- Gonzalez, R., A. Butkovic and S. F. Elena, “From foes to friends: viral infections expand the limits of host phenotypic plasticity”, *Adv. Virus Res.* **106** (2020), pp. 85–121.
- Gorbalenya, A. E., L. Enjuanes, J. Ziebuhr and E. J. Snijder, “Nidovirales: evolving the largest RNA virus genome”, *Virus Res.* **117** (2006), no. 1, pp. 17–37.
- Gorbalenya, A. E. and E. V. Koonin, “Viral proteins containing the purine NTP-binding sequence pattern”, *Nucleic Acids Res.* **17** (1989), no. 21, pp. 8413–8440.
- Gorlas, A., E. V. Koonin, N. Bienvenu, D. Prieur and C. Geslin, “TPV1, the first virus isolated from the hyperthermophilic genus *Thermococcus*”, *Environ. Microbiol.* **14** (2012), no. 2, pp. 503–516.
- Gozzi, K., N. T. Tran, J. W. Modell, T. B. K. Le and M. T. Laub, “Prophage-like gene transfer agents promote *Caulobacter crescentus* survival and DNA repair during stationary phase”, *PLoS Biol.* **20** (2022), no. 11, article no. e3001790.
- Grimm, C., H. S. Hillen, K. Bedenk, et al., “Structural basis of poxvirus transcription: vaccinia RNA polymerase complexes”, *Cell* **179** (2019), no. 7, pp. 1537–1550.
- Guardado-Calvo, P. and F. A. Rey, “The viral class II membrane fusion machinery: divergent evolution from an ancestral heterodimer”, *Viruses* **13** (2021), no. 12, article no. 2368.
- Guinet, B., M. Leobold, E. A. Herniou, et al., “A novel and diverse family of filamentous DNA viruses associated with parasitic wasps”, *Virus Evol.* **10** (2024), no. 1, article no. veae022.
- Han, Z., W. Yuan, H. Xiao, et al., “Structural insights into a spindle-shaped archaeal virus with a sevenfold symmetrical tail”, *Proc. Natl. Acad. Sci. USA* **119** (2022), no. 31, article no. e2119439119.
- Harrison, E. and M. A. Brockhurst, “Ecological and evolutionary benefits of temperate phage: what does or doesn’t kill you makes you stronger”, *Bioessays* **39** (2017), no. 12, article no. 1700112.
- Hatfull, G. F., “Actinobacteriophages: genomics, dynamics, and applications”, *Annu. Rev. Virol.* **7** (2020), no. 1, pp. 37–61.
- Hatfull, G. F., “Mycobacteriophages: from Petri dish to patient”, *PLoS Pathog.* **18** (2022), no. 7, article no. e1010602.
- Hayes, V. M., J. T. Zhang, M. A. Katz, Y. Li, B. Kocsis, D. M. Brinkley, N. Jia and A. J. Meeske, “RNA-mediated CRISPR-Cas13 inhibition through crRNA structural mimicry”, *Science* **388** (2025), no. 6745, pp. 387–391.
- Hendrix, R. W., “Bacteriophage genomics”, *Curr. Opin. Microbiol.* **6** (2003), no. 5, pp. 506–511.
- Hepojoki, J., U. Hetzel, S. Paraskevopoulou, et al., *Create one new realm (Ribozyviria) including one new family (Kolmioviridae) including genus Deltavirus and seven new genera for a total of 15 species*, Technical report, International Committee on Taxonomy of Viruses Taxonomic Proposal, 2021. Online at <https://ictv.global/ictv/proposals/2020.012D.R.Ribozyviria.zip> (accessed on April 27, 2026).
- Hillen, H. S., J. Bartuli, C. Grimm, C. Dienemann, K. Bedenk, A. A. Szalay, U. Fischer and P. Cramer, “Structural basis of poxvirus transcription: transcribing and capping vaccinia complexes”, *Cell* **179** (2019), no. 7, pp. 1525–1536.
- Hou, X., Y. He, P. Fang, et al., “Using artificial intelligence to document the hidden RNA virosphere”, *Cell* **187** (2024), no. 24, pp. 6929–6942.
- Humolli, D., D. Piel, E. Maffei, et al., “Completing the BASEL phage collection to unlock hidden diversity for systematic exploration of phage-host interactions”, *PLoS Biol.* **23** (2025), no. 4, article no. e3003063.
- Ilyina, T. V. and E. V. Koonin, “Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaeobacteria”, *Nucleic Acids Res.* **20** (1992), no. 13, pp. 3279–3285.
- International Committee on Taxonomy of Viruses Executive Committee, “The new scope of virus taxonomy: partitioning the virosphere into 15 hierarchical ranks”, *Nat. Microbiol.* **5** (2020), no. 5, pp. 668–674.
- Iranzo, J., E. V. Koonin, D. Prangishvili and M. Krupovic, “Bipartite network analysis of the archaeal virosphere: evolutionary connections between viruses and capsidless mobile elements”, *J. Virol.* **90** (2016), no. 24, pp. 11043–11055.
- Iranzo, J., M. Krupovic and E. V. Koonin, “The double-stranded DNA virosphere as a modular hierarchical network of gene sharing”, *mBio* **7** (2016), no. 4, article no. e00978-16.
- Iyer, L. M., S. Abhiman and L. Aravind, “A new family of polymerases related to superfamily A DNA polymerases and T7-like DNA-dependent RNA polymerases”, *Biol. Direct* **3** (2008), article no. 39.
- Iyer, L. M., E. V. Koonin, D. D. Leipe and L. Aravind, “Origin and evolution of the archaeo-eukaryotic primase superfamily and related palm-domain proteins: structural insights and new members”, *Nucleic Acids Res.* **33** (2005), no. 12, pp. 3875–3896.
- Iyer, L. M., K. S. Makarova, E. V. Koonin and L. Aravind, “Comparative genomics of the FtsK-HerA superfamily of pumping ATPases: implications for the origins of chromosome segregation, cell division and viral capsid packaging”, *Nucleic Acids Res.* **32** (2004), no. 17, pp. 5260–5279.
- Jeong, D. E., S. Sundrani, R. N. Hall, M. Krupovic, E. V. Koonin and A. Z. Fire, “DNA polymerase diversity reveals multiple incur-

- sions of polintons during nematode evolution”, *Mol. Biol. Evol.* **40** (2023), no. 12, article no. msad274.
- Johnstone, B. A., J. M. Hardy, J. Ha, et al., “The nucleocapsid architecture and structural atlas of the prototype baculovirus define the hallmarks of a new viral realm”, *Sci. Adv.* **10** (2024), no. 51, article no. eado2631.
- Joyce, G. F., “The antiquity of RNA-based evolution”, *Nature* **418** (2002), no. 6894, pp. 214–221.
- Kaczorowska, J. and L. van der Hoek, “Human anelloviruses: diverse, omnipresent and commensal members of the virome”, *FEMS Microbiol. Rev.* **44** (2020), no. 3, pp. 305–313.
- Kapitonov, V. V. and J. Jurka, “Self-synthesizing DNA transposons in eukaryotes”, *Proc. Natl. Acad. Sci. USA* **103** (2006), no. 12, pp. 4540–4545.
- Kasson, P., F. DiMaio, X. Yu, S. Lucas-Staat, M. Krupovic, S. Schouten, D. Prangishvili and E. H. Egelman, “Model for a novel membrane envelope in a filamentous hyperthermophilic virus”, *Elife* **6** (2017), article no. e26268.
- Kazlauskas, D., A. Dayaram, S. Kraberger, S. Goldstien, A. Varsani and M. Krupovic, “Evolutionary history of ssDNA bacilladnaviruses features horizontal acquisition of the capsid gene from ssRNA nodaviruses”, *Virology* **504** (2017), pp. 114–121.
- Kazlauskas, D., M. Krupovic, J. Guglielmini, P. Forterre and C. Venclovas, “Diversity and evolution of B-family DNA polymerases”, *Nucleic Acids Res.* **48** (2020), no. 18, pp. 10142–10156.
- Kazlauskas, D., M. Krupovic and C. Venclovas, “The logic of DNA replication in double-stranded DNA viruses: insights from global analysis of viral genomes”, *Nucleic Acids Res.* **44** (2016), no. 10, pp. 4551–4564.
- Kazlauskas, D., A. Varsani, E. V. Koonin and M. Krupovic, “Multiple origins of prokaryotic and eukaryotic single-stranded DNA viruses from bacterial and archaeal plasmids”, *Nat. Commun.* **10** (2019), no. 1, article no. 3425.
- Khalfi, P., P. T. Kennedy, K. Majzoub and T. Asselah, “Hepatitis D virus: improving virological knowledge to develop new treatments”, *Antiviral Res.* **209** (2023), article no. 105461.
- Kielian, M. and E. A. Rey, “Virus membrane-fusion proteins: more than one way to make a hairpin”, *Nat. Rev. Microbiol.* **4** (2006), no. 1, pp. 67–76.
- Kim, J. G., S. J. Kim, V. Cvirkaite-Krupovic, et al., “Spindle-shaped viruses infect marine ammonia-oxidizing thaumarchaea”, *Proc. Natl. Acad. Sci. USA* **116** (2019), no. 31, pp. 15645–15650.
- Kirchberger, P. C., Z. A. Martinez, L. J. Luker and H. Ochman, “Defensive hypervariable regions confer superinfection exclusion in microviruses”, *Proc. Natl. Acad. Sci. USA* **118** (2021), no. 28, article no. e2102786118.
- Kirchberger, P. C., Z. A. Martinez and H. Ochman, “Organizing the global diversity of microviruses”, *mBio* **13** (2022), no. 3, article no. e0058822.
- Kirchberger, P. C. and H. Ochman, “Resurrection of a global, metagenomically defined gokushovirus”, *Elife* **9** (2020), article no. e51599.
- Kirchberger, P. C. and H. Ochman, “Microviruses: a world beyond phiX174”, *Annu. Rev. Virol.* **10** (2023), no. 1, pp. 99–118.
- Knowles, B., B. Bailey, L. Boling, et al., “Variability and host density independence in inductions-based estimates of environmental lysogeny”, *Nat. Microbiol.* **2** (2017), article no. 17064.
- Kogay, R. and O. Zhaxybayeva, “Co-evolution of gene transfer agents and their alphaproteobacterial hosts”, *J. Bacteriol.* **206** (2024), no. 2, article no. e0039823.
- Kojima, K. K., “Troyka represents a unique lineage of virus-like retroelements”, *J. Gen. Virol.* **106** (2025), no. 9, article no. 002143.
- Koonin, E. V., “Comparative genomics, minimal gene-sets and the last universal common ancestor”, *Nat. Rev. Microbiol.* **1** (2003), no. 2, pp. 127–136.
- Koonin, E. V. and V. V. Dolja, “Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences”, *Crit. Rev. Biochem. Mol. Biol.* **28** (1993), no. 5, pp. 375–430.
- Koonin, E. V. and V. V. Dolja, “A virocentric perspective on the evolution of life”, *Curr. Opin. Virol.* **3** (2013), no. 5, pp. 546–557.
- Koonin, E. V. and V. V. Dolja, “Virus world as an evolutionary network of viruses and capsidless selfish elements”, *Microbiol. Mol. Biol. Rev.* **78** (2014), no. 2, pp. 278–303.
- Koonin, E. V., V. V. Dolja and M. Krupovic, “The healthy human virome: from virus–host symbiosis to disease”, *Curr. Opin. Virol.* **47** (2021), pp. 86–94.
- Koonin, E. V., V. V. Dolja and M. Krupovic, “The logic of virus evolution”, *Cell Host Microb.* **30** (2022), no. 7, pp. 917–929.
- Koonin, E. V., V. V. Dolja, M. Krupovic and J. H. Kuhn, “Viruses defined by the position of the virosphere within the replicator space”, *Microbiol. Mol. Biol. Rev.* **85** (2021), no. 4, article no. e0019320.
- Koonin, E. V., V. V. Dolja, M. Krupovic, A. Varsani, Y. I. Wolf, N. Yutin, F. M. Zerbini and J. H. Kuhn, “Global organization and proposed megataxonomy of the virus world”, *Microbiol. Mol. Biol. Rev.* **84** (2020), no. 2, article no. e00061-19.
- Koonin, E. V., M. G. Fischer, J. H. Kuhn and M. Krupovic, “The polinton-like supergroup of viruses: evolution, molecular biology, and taxonomy”, *Microbiol. Mol. Biol. Rev.* **88** (2024), no. 3, article no. e0008623.
- Koonin, E. V., M. Krupovic and V. I. Agol, “The Baltimore classification of viruses 50 years later: how does it stand in the light of virus evolution?”, *Microbiol. Mol. Biol. Rev.* **85** (2021), no. 3, article no. e0005321.
- Koonin, E. V., M. Krupovic and V. V. Dolja, “The global virome: how much diversity and how many independent origins?”, *Environ. Microbiol.* **25** (2023), no. 1, pp. 40–44.
- Koonin, E. V., M. Krupovic, S. Ishino and Y. Ishino, “The replication machinery of LUCA: common origin of DNA replication and transcription”, *BMC Biol.* **18** (2020), no. 1, article no. 61.
- Koonin, E. V., J. H. Kuhn, V. V. Dolja and M. Krupovic, “Megataxonomy and global ecology of the virosphere”, *ISME J.* **18** (2024), no. 1, article no. wrad042.
- Koonin, E. V., K. S. Makarova and Y. I. Wolf, “Evolution of microbial genomics: conceptual shifts over a quarter century”, *Trends Microbiol.* **29** (2021), no. 7, pp. 582–592.
- Koonin, E. V., K. S. Makarova, Y. I. Wolf and M. Krupovic, “Evolutionary entanglement of mobile genetic elements and host defence systems: guns for hire”, *Nat. Rev. Genet.* **21** (2020), no. 2, pp. 119–131.
- Koonin, E. V., T. G. Senkevich and V. V. Dolja, “The ancient virus world and evolution of cells”, *Biol. Direct* **1** (2006), article no. 29.

- Koonin, E. V., Y. I. Wolf and M. I. Katsnelson, "Inevitability of the emergence and persistence of genetic parasites caused by evolutionary instability of parasite-free states", *Biol. Direct* **12** (2017), no. 1, article no. 31.
- Koonin, E. V. and N. Yutin, "Multiple evolutionary origins of giant viruses", *F1000Res* **7** (2018), article no. 1840.
- Koonin, E. V. and N. Yutin, "Evolution of the large nucleocytoplasmic DNA viruses of eukaryotes and convergent origins of viral gigantism", *Adv. Virus Res.* **103** (2019), pp. 167–202.
- Koslova, A., T. Hackl, F. Bade, A. Sanchez Kasikovic, K. Barenhoff, F. Schimm, U. Mersdorf and M. G. Fischer, "Endogenous virophages are active and mitigate giant virus infection in the marine protist *Cafeteria burkhardae*", *Proc. Natl. Acad. Sci. USA* **121** (2024), no. 11, article no. e2314606121.
- Krabberger, S., T. Opriessnig, F. Maggi, et al., "ICTV virus taxonomy profile: *Anelloviridae* 2026", *J. Gen. Virol.* **107** (2026), no. 3, article no. 002222.
- Krishnamurthy, S. R. and D. Wang, "Extensive conservation of prokaryotic ribosomal binding sites in known and novel picobirnaviruses", *Virology* **516** (2018), pp. 108–114.
- Krupovic, M. and D. H. Bamford, "Putative prophages related to lytic tailless marine dsDNA phage PM2 are widespread in the genomes of aquatic bacteria", *BMC Genom.* **8** (2007), article no. 236.
- Krupovic, M. and D. H. Bamford, "Virus evolution: how far does the double beta-barrel viral lineage extend?", *Nat. Rev. Microbiol.* **6** (2008), no. 12, pp. 941–948.
- Krupovic, M. and D. H. Bamford, "Order to the viral universe", *J. Virol.* **84** (2010), no. 24, pp. 12476–12479.
- Krupovic, M., D. H. Bamford and E. V. Koonin, "Conservation of major and minor jelly-roll capsid proteins in Polinton (Maverick) transposons suggests that they are bona fide viruses", *Biol. Direct* **9** (2014), article no. 6.
- Krupovic, M., J. Blomberg, J. M. Coffin, et al., "Ortervirales: new virus order unifying five families of reverse-transcribing viruses", *J. Virol.* **92** (2018), no. 12, article no. e00515-18.
- Krupovic, M., V. V. Dolja and E. V. Koonin, "Origin of viruses: primordial replicators recruiting capsids from hosts", *Nat. Rev. Microbiol.* **17** (2019), pp. 449–458.
- Krupovic, M., V. V. Dolja and E. V. Koonin, "The LUCA and its complex virome", *Nat. Rev. Microbiol.* **18** (2020), no. 11, pp. 661–670.
- Krupovic, M., V. V. Dolja and E. V. Koonin, "The virome of the last eukaryotic common ancestor and eukaryogenesis", *Nat. Microbiol.* **8** (2023), no. 6, pp. 1008–1017.
- Krupovic, M. and P. Forterre, "Microviridae goes temperate: microvirus-related proviruses reside in the genomes of Bacteroidetes", *PLoS One* **6** (2011), no. 5, article no. e19893.
- Krupovic, M. and P. Forterre, "Single-stranded DNA viruses employ a variety of mechanisms for integration into the host genomes", *Ann N.Y. Acad. Sci.* **1341** (2015), pp. 41–53.
- Krupovic, M. and ICTV Report C, "ICTV virus taxonomy profile: *Plasmaviridae*", *J. Gen. Virol.* **99** (2018), no. 5, pp. 617–618.
- Krupovic, M. and E. V. Koonin, "Evolution of eukaryotic single-stranded DNA viruses of the Bidnaviridae family from genes of four other groups of widely different viruses", *Sci. Rep.* **4** (2014), article no. 5347.
- Krupovic, M. and E. V. Koonin, "Polintons: a hotbed of eukaryotic virus, transposon and plasmid evolution", *Nat. Rev. Microbiol.* **13** (2015), no. 2, pp. 105–115.
- Krupovic, M. and E. V. Koonin, "Homologous capsid proteins testify to the common ancestry of retroviruses, caulimoviruses, pseudoviruses and metaviruses", *J. Virol.* **91** (2017a), no. 12, article no. e00210-17.
- Krupovic, M. and E. V. Koonin, "Multiple origins of viral capsid proteins from cellular ancestors", *Proc. Natl. Acad. Sci. USA* **114** (2017b), no. 12, E2401–E2410.
- Krupovic, M., J. H. Kuhn, M. G. Fischer and E. V. Koonin, "Natural history of eukaryotic DNA viruses with double jelly-roll major capsid proteins", *Proc. Natl. Acad. Sci. USA* **121** (2024), no. 23, article no. e2405771121.
- Krupovic, M., J. H. Kuhn, F. Wang, D. P. Baquero, V. V. Dolja, E. H. Egelman, D. Prangishvili and E. V. Koonin, "Adnaviria: a new realm for archaeal filamentous viruses with linear A-form double-stranded DNA genomes", *J. Virol.* **95** (2021), no. 15, article no. e0067321.
- Krupovic, M., K. S. Makarova and E. V. Koonin, "Cellular homologs of the double jelly-roll major capsid proteins clarify the origins of an ancient virus kingdom", *Proc. Natl. Acad. Sci. USA* **119** (2022), no. 5, article no. e2120620119.
- Krupovic, M., E. R. Quemin, D. H. Bamford, P. Forterre and D. Prangishvili, "Unification of the globally distributed spindle-shaped viruses of the Archaea", *J. Virol.* **88** (2014), no. 4, pp. 2354–2358.
- Krupovic, M. and A. Varsani, "Naryaviridae, Nenyaviridae, and Vilyaviridae: three new families of single-stranded DNA viruses in the phylum Cressdnaviricota", *Arch. Virol.* **167** (2022), no. 12, pp. 2907–2921.
- Krupovic, M., A. Varsani, D. Kazlauskas, et al., "Cressdnaviricota: a virus phylum unifying seven families of rep-encoding viruses with single-stranded, circular DNA genomes", *J. Virol.* **94** (2020), no. 12, article no. e00582-20.
- Krupovic, M., F. Wang, V. Cvirkaite-Krupovic, J. H. Kuhn, V. V. Dolja, D. Prangishvili, E. H. Egelman and E. V. Koonin, "ICTV virus taxonomy profile: Adnaviria 2025", *J. Gen. Virol.* **106** (2025), no. 3, article no. 002091.
- Krupovic, M., N. Yutin and E. V. Koonin, "Fusion of a superfamily 1 helicase and an inactivated DNA polymerase is a signature of common evolutionary history of Polintons, polinton-like viruses, Tlr1 transposons and transpovirons", *Virus Evol.* **2** (2016), no. 1, article no. vew019.
- Krupovic, M., N. Yutin and E. V. Koonin, "Evolution of a major virion protein of the giant pandoraviruses from an inactivated bacterial glycoside hydrolase", *Virus Evol.* **6** (2020), no. 2, article no. veaa059.
- Kuhn, J. H., A. Babiain, L. M. Bergner, et al., "ICTV virus taxonomy profile: *Kolmioviridae* 2024", *J. Gen. Virol.* **105** (2024), no. 2, article no. 001963.
- Kuhn, J. H., L. Botella, M. de la Pena, et al., "Ambiviricota, a novel ribovirion phylum for viruses with viroid-like properties", *J. Virol.* **98** (2024), no. 7, article no. e0083124.
- Kuhn, J. H. and E. V. Koonin, "Viriforms—a new category of classifiable virus-derived genetic elements", *Biomolecules* **13** (2023), no. 2, article no. 289.

- La Scola, B., C. Desnues, I. Pagnier, et al., "The virophage as a unique parasite of the giant mimivirus", *Nature* **455** (2008), no. 7209, pp. 100–104.
- Laanto, E., S. Mantynen, L. De Colibus, et al., "Virus found in a boreal lake links ssDNA and dsDNA viruses", *Proc. Natl. Acad. Sci. USA* **114** (2017), no. 31, pp. 8378–8383.
- Lang, A. S., O. Zhaxybayeva and J. T. Beatty, "Gene transfer agents: phage-like elements of genetic exchange", *Nat. Rev. Microbiol.* **10** (2012), no. 7, pp. 472–482.
- Laso-Perez, R., F. Wu, A. Cremiere, D. R. Speth, J. S. Magyar, K. Zhao, M. Krupovic and V. J. Orphan, "Evolutionary diversification of methanotrophic ANME-1 archaea and their expansive virome", *Nat. Microbiol.* **8** (2023), no. 2, pp. 231–245.
- Lee, B. D., U. Neri, S. Roux, et al., "Mining metatranscriptomes reveals a vast world of viroid-like circular RNAs", *Cell* **186** (2023), no. 3, pp. 646–661.
- Lee, H., A. J. Baxter, C. M. Bator, B. A. Fane and S. L. Hafenstein, "Cryo-EM structure of Gokushovirus PhiEC6098 reveals a novel capsid architecture for a single-scaffolding protein, microvirus assembly system", *J. Virol.* **96** (2022), no. 21, article no. e0099022.
- Lee, J., J. Lee, E. J. Sohn, A. Tagliatalata, R. J. O'Sullivan, A. Ciccio and J. Min, "Extrachromosomal telomere DNA derived from excessive strand displacements", *Proc. Natl. Acad. Sci. USA* **121** (2024), no. 19, article no. e2318438121.
- Lee, J., E. J. Sohn, J. Lee, A. Y. Cheng, A. Tagliatalata, A. Ciccio and J. Min, "Distinct mechanisms underlying extrachromosomal telomere DNA generation in ALT cancers", *Nucleic Acids Res.* **53** (2025), no. 15, article no. gkaf771.
- Legendre, M., J. M. Alempic, N. Philippe, et al., "Pandoravirus Celtis illustrates the microevolution processes at work in the giant Pandoraviridae genomes", *Front. Microbiol.* **10** (2019), article no. 430.
- Legendre, M., J. Bartoli, L. Shmakova, et al., "Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology", *Proc. Natl. Acad. Sci. USA* **111** (2014), no. 11, pp. 4274–4279.
- Legendre, M., E. Fabre, O. Poirot, et al., "Diversity and evolution of the emerging Pandoraviridae family", *Nat. Commun.* **9** (2018), no. 1, article no. 2285.
- Li, D., R. Zhao, W. Lilyestrom, et al., "Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen", *Nature* **423** (2003), no. 6939, pp. 512–518.
- Li, S., T. Xu, X. Meng, et al., "Ocr-mediated suppression of BrxX unveils a phage counter-defense mechanism", *Nucleic Acids Res.* **52** (2024), no. 14, pp. 8580–8594.
- Li, W. X. and S. W. Ding, "Mammalian viral suppressors of RNA interference", *Trends Biochem. Sci.* **47** (2022), no. 11, pp. 978–988.
- Liou, S. H., R. Boggavarapu, N. R. Cohen, et al., "Structure of anellovirus-like particles reveal a mechanism for immune evasion", *Nat. Commun.* **15** (2024), no. 1, article no. 7219.
- Liu, Y., T. A. Demina, S. Roux, et al., "Diversity, taxonomy, and evolution of archaeal viruses of the class Caudoviricetes", *PLoS Biol.* **19** (2021), no. 11, article no. e3001442.
- Liu, Y., M. Dyall-Smith and H. M. Oksanen, "ICTV virus taxonomy profile: *Pleolipoviridae* 2022", *J. Gen. Virol.* **103** (2022), no. 11, article no. 001793.
- Liu, Y., S. Ishino, Y. Ishino, G. Pehau-Arnaudet, M. Krupovic and D. Prangishvili, "A novel type of polyhedral viruses infecting hyperthermophilic archaea", *J. Virol.* **91** (2017), no. 13, article no. e00589-17.
- Liu, Y., T. Osinski, F. Wang, M. Krupovic, S. Schouten, P. Kasson, D. Prangishvili and E. H. Egelman, "Structural conservation in a membrane-enveloped filamentous virus infecting a hyperthermophilic acidophile", *Nat. Commun.* **9** (2018), no. 1, article no. 3360.
- Lombard, J., P. Lopez-Garcia and D. Moreira, "The early evolution of lipid membranes and the three domains of life", *Nat. Rev. Microbiol.* **10** (2012), no. 7, pp. 507–515.
- Lopez-Garcia, P., "The place of viruses in biology in light of the metabolism- versus-replication-first debate", *Hist. Philos. Life Sci.* **34** (2012), no. 3, pp. 391–406.
- Lopez-Garcia, P., A. Gutierrez-Preciado, M. Krupovic, et al., "Metagenome-derived virus-microbe ratios across ecosystems", *ISME J.* **17** (2023), no. 10, pp. 1552–1563.
- Lopez-Garcia, P. and D. Moreira, "The Syntrophy hypothesis for the origin of eukaryotes revisited", *Nat. Microbiol.* **5** (2020), no. 5, pp. 655–667.
- Lopez-Simon, J., M. de la Pena and M. Martinez-Garcia, "Viroid-like "obelisk" agents are widespread in the ocean and exceed the abundance of RNA viruses in the prokaryotic fraction", *ISME J.* **19** (2025), no. 1, article no. wraf033.
- Lu, C., Y. Wu, Z. Zhang, et al., "Estimated size of the total genome and protein space of viruses", *mSphere* **10** (2025), no. 3, article no. e0068324.
- Lund, M. C., A. Hopkins, A. Dayaram, et al., "Diverse microviruses circulating in invertebrates within a lake ecosystem", *J. Gen. Virol.* **105** (2024), no. 11, article no. 002049.
- Luque, A., S. Benler, D. Y. Lee, C. Brown and S. White, "The missing tailed phages: prediction of small capsid candidates", *Microorganisms* **8** (2020), no. 12, article no. 1944.
- Luria, S. E. and J. Darnell, *General Virology*, John Wiley: New York, 1967.
- Lwoff, A., "Lysogeny", *Bacteriol. Rev.* **17** (1953), no. 4, pp. 269–337.
- Makarova, K. S., Y. I. Wolf, P. Forterre, D. Prangishvili, M. Krupovic and E. V. Koonin, "Dark matter in archaeal genomes: a rich source of novel mobile elements, defense systems and secretory complexes", *Extremophiles* **18** (2014), no. 5, pp. 877–893.
- Malik, H. S. and T. H. Eickbush, "Phylogenetic analysis of ribonuclease H domains suggests a late, chimeric origin of LTR retrotransposable elements and retroviruses", *Genome Res.* **11** (2001), no. 7, pp. 1187–1197.
- Maynard Smith, J. and E. Szathmary, *The Major Transitions in Evolution*, W. H. Freeman: New York, 1995.
- McKellar, J., A. Fouillen, S. Lyonnais, et al., "Deltaviruses spread through a viral Trojan Horse", *Cell* **189** (2026), pp. 1–14.
- Medawar, P. B., *Immunological tolerance*, Nobel Lecture, 1960. Online at <https://www.nobelprize.org/prizes/medicine/1960/medawar/lecture/> (accessed on April 27, 2026).
- Medvedeva, S., G. Borrel, M. Krupovic and S. Gribaldo, "A compendium of viruses from methanogenic archaea reveals their diversity and adaptations to the gut environment", *Nat. Microbiol.* **8** (2023), no. 11, pp. 2170–2182.
- Medvedeva, S., U. Guyet, E. Pelletier, et al., "Widespread and intron-rich mirusviruses are predicted to reproduce in nuclei

- of unicellular eukaryotes”, *Nat. Microbiol.* **11** (2026), no. 1, pp. 228–239.
- Medvedeva, S., Y. Liu, E. V. Koonin, K. Severinov, D. Prangishvili and M. Krupovic, “Virus-borne mini-CRISPR arrays are involved in interviral conflicts”, *Nat. Commun.* **10** (2019), no. 1, article no. 5204.
- Medvedeva, S., J. Sun, N. Yutin, E. V. Koonin, T. Nunoura, C. Rinke and M. Krupovic, “Three families of Asgard archaeal viruses identified in metagenome-assembled genomes”, *Nat. Microbiol.* **7** (2022), no. 7, pp. 962–973.
- Mietzsch, M., A. Bennett and R. McKenna, “Structural capsidomics of single-stranded DNA viruses”, *Viruses* **17** (2025), no. 3, article no. 333.
- Mietzsch, M., S. Kailasan, A. Bennett, P. Chipman, B. Fane, J. T. Huisken, I. N. Clarke and R. McKenna, “The structure of Spiroplasma virus 4: exploring the capsid diversity of the microviridae”, *Viruses* **16** (2024), no. 7, article no. 1103.
- Mizuno, C. M., C. Guyomar, S. Roux, et al., “Numerous cultivated and uncultivated viruses encode ribosomal proteins”, *Nat. Commun.* **10** (2019), no. 1, article no. 752.
- Moody, E. R. R., S. Alvarez-Caretero, T. A. Mahendrarajah, et al., “The nature of the last universal common ancestor and its impact on the early Earth system”, *Nat. Ecol. Evol.* **8** (2024), no. 9, pp. 1654–1666.
- Moreira, D. and C. Brochier-Armanet, “Giant viruses, giant chimeras: the multiple evolutionary histories of Mimivirus genes”, *BMC Evol. Biol.* **8** (2008), article no. 12.
- Moreira, D. and P. Lopez-Garcia, “Symbiosis between methanogenic archaea and delta-proteobacteria as the origin of eukaryotes: the syntrophic hypothesis”, *J. Mol. Evol.* **47** (1998), no. 5, pp. 517–530.
- Moreira, D. and P. Lopez-Garcia, “Ten reasons to exclude viruses from the tree of life”, *Nat. Rev. Microbiol.* **7** (2009), no. 4, pp. 306–311.
- Moreira, D. and P. Lopez-Garcia, “Evolution of viruses and cells: do we need a fourth domain of life to explain the origin of eukaryotes?”, *Philos. Trans. R. Soc. Lond. B* **370** (2015), no. 1678, article no. 20140327.
- Munke, A., K. Kimura, Y. Tomaru, H. Wang, K. Yoshida, S. Mito, Y. Hongo and K. Okamoto, “Primordial capsid and spooled ss-DNA genome structures unravel ancestral events of eukaryotic viruses”, *mBio* **13** (2022), no. 4, article no. e0015622.
- Mushegian, A. R., “Are there 10(31) virus particles on Earth, or more, or fewer?”, *J. Bacteriol.* **202** (2020), no. 9, article no. e00052-00020.
- Mushegian, A. R. and S. F. Elena, “Evolution of plant virus movement proteins from the 30K superfamily and of their homologs integrated in plant genomes”, *Virology* **476** (2015), pp. 304–315.
- Nasir, A., K. M. Kim and G. Caetano-Anolles, “Giant viruses coexisted with the cellular ancestors and represent a distinct supergroup along with superkingdoms Archaea, Bacteria and Eukarya”, *BMC Evol. Biol.* **12** (2012), article no. 156.
- Nasir, A., K. M. Kim and G. Caetano-Anolles, “Phylogenetic tracings of proteome size support the gradual accretion of protein structural domains and the early origin of viruses from primordial cells”, *Front. Microbiol.* **8** (2017), article no. 1178.
- Nelson, C. A., M. L. Epperson, S. Singh, J. I. Elliott and D. H. Fremont, “Structural conservation and functional diversity of the poxvirus immune evasion (PIE) domain superfamily”, *Viruses* **7** (2015), no. 9, pp. 4878–4898.
- Neri, U., Y. I. Wolf, S. Roux, et al., “Expansion of the global RNA virome reveals diverse clades of bacteriophages”, *Cell* **185** (2022), no. 21, pp. 4023–4037.
- Neuman, B. W., A. Smart, O. Gilmer, et al., “Giant RNA genomes: roles of host, translation elongation, genome architecture, and proteome in nidoviruses”, *Proc. Natl. Acad. Sci. USA* **122** (2025), no. 7, article no. e2413675122.
- Nissimov, J. I., D. Talmy, L. Haramaty, et al., “Biochemical diversity of glycosphingolipid biosynthesis as a driver of Coccolithovirus competitive ecology”, *Environ. Microbiol.* **21** (2019), no. 6, pp. 2182–2197.
- Oke, M., M. Kerou, H. Liu, X. Peng, R. A. Garrett, D. Prangishvili, J. H. Naismith and M. F. White, “A dimeric Rep protein initiates replication of a linear archaeal virus genome: implications for the Rep mechanism and viral replication”, *J. Virol.* **85** (2011), no. 2, pp. 925–931.
- Olo Ndela, E., S. Roux, C. Henke, A. Sczyrba, T. Sime Ngando, A. Varsani and F. Enault, “Reekeek- and roodoooviruses, two different Microviridae clades constituted by the smallest DNA phages”, *Virus Evol.* **9** (2023), no. 1, article no. veac123.
- Paraskevopoulou, S., F. Pirzer, N. Goldmann, et al., “Mammalian deltavirus without hepadnavirus coinfection in the neotropical rodent *Proechimys semispinosus*”, *Proc. Natl. Acad. Sci. USA* **117** (2020), no. 30, pp. 17977–17983.
- Pastuzyn, E. D., C. E. Day, R. B. Kearns, et al., “The neuronal gene arc encodes a repurposed retrotransposon gag protein that mediates intercellular RNA transfer”, *Cell* **172** (2018), no. 1–2, pp. 275–288.
- Pawluk, A., A. R. Davidson and K. L. Maxwell, “Anti-CRISPR: discovery, mechanism and function”, *Nat. Rev. Microbiol.* **16** (2018), no. 1, pp. 12–17.
- Perez-Vargas, J., R. Pereira de Oliveira, S. Jacquet, D. Pontier, F. L. Cosset and N. Freitas, “HDV-like viruses”, *Viruses* **13** (2021), no. 7, article no. 1207.
- Perflyev, A., A. Gaede, S. Hooton, et al., “Large-scale analysis of bacterial genomes reveals thousands of lytic phages”, *Nat. Microbiol.* **11** (2026), no. 1, pp. 42–52.
- Petersen, J. M., A. Bezier, J. M. Drezen and M. M. van Oers, “The naked truth: an updated review on nudiviruses and their relationship to bracoviruses and baculoviruses”, *J. Invertebr. Pathol.* **189** (2022), article no. 107718.
- Philippe, N., M. Legendre, G. Doutre, et al., “Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes”, *Science* **341** (2013), no. 6143, pp. 281–286.
- Philosof, A., N. Yutin, J. Flores-Urbe, I. Sharon, E. V. Koonin and O. Beja, “Novel abundant oceanic viruses of uncultured marine group II euryarchaeota”, *Curr. Biol.* **27** (2017), no. 9, pp. 1362–1368.
- Pidre, M. L., P. N. Arrias, L. C. Amoros Morales and V. Romanowski, “The magic staff: a comprehensive overview of baculovirus-based technologies applied to human and animal health”, *Viruses* **15** (2022), no. 1, article no. 80.
- Prangishvili, D., Y. Liu, M. Krupovic and ICTV Report C, “ICTV virus taxonomy profile: *Portogloboviridae*”, *J. Gen. Virol.* **102** (2021), no. 6, article no. 001605.

- Pritham, E. J., T. Putliwala and C. Feschotte, “Mavericks, a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses”, *Gene* **390** (2007), no. 1–2, pp. 3–17.
- Puxty, R. J., A. D. Millard, D. J. Evans and D. J. Scanlan, “Shedding new light on viral photosynthesis”, *Photosynth. Res.* **126** (2015), no. 1, pp. 71–97.
- Queiroz, V. F., J. Carvalho, F. G. de Souza, et al., “Analysis of the genomic features and evolutionary history of pithovirus-like isolates reveals two major divergent groups of viruses”, *J. Virol.* **97** (2023), no. 7, article no. e0041123.
- Rakonjac, J., V. A. M. Gold, R. I. Leon-Quezada and C. H. Davenport, “Structure, biology, and applications of filamentous bacteriophages”, *Cold Spring Harb. Protoc.* **2024** (2024), no. 8, article no. pdb.over107754.
- Rambo, I. M., M. V. Langwig, P. Leao, V. De Anda and B. J. Baker, “Genomes of six viruses that infect Asgard archaea from deep-sea sediments”, *Nat. Microbiol.* **7** (2022), no. 7, pp. 953–961.
- Raoult, D., S. Audic, C. Robert, et al., “The 1.2-megabase genome sequence of Mimivirus”, *Science* **306** (2004), no. 5700, pp. 1344–1350.
- Raoult, D. and P. Forterre, “Redefining viruses: lessons from Mimivirus”, *Nat. Rev. Microbiol.* **6** (2008), no. 4, pp. 315–319.
- Rinke, C., P. Schwientek, A. Sczyrba, et al., “Insights into the phylogeny and coding potential of microbial dark matter”, *Nature* **499** (2013), no. 7459, pp. 431–437.
- Rissanen, I., J. M. Grimes, A. Pawlowski, S. Mantynen, K. Harlos, J. K. Bamford and D. I. Stuart, “Bacteriophage P23-77 capsid protein structures reveal the archetype of an ancient branch from a major virus lineage”, *Structure* **21** (2013), no. 5, pp. 718–726.
- Rochman, N. D., Y. I. Wolf and E. V. Koonin, “Molecular adaptations during viral epidemics”, *EMBO Rep.* **23** (2022), no. 8, article no. e55393.
- Rodrigues, R. A. L., T. S. Arantes, G. P. Oliveira, L. K. Dos Santos Silva and J. S. Abrahao, “The complex nature of Tupanviruses”, *Adv. Virus Res.* **103** (2019), pp. 135–166.
- Rodrigues, R. A. L., L. C. F. da Silva and J. S. Abrahao, “Translating the language of giants: translation-related genes as a major contribution of giant viruses to the virosphere”, *Arch. Virol.* **165** (2020), no. 6, pp. 1267–1278.
- Roine, E., P. Kukkaro, L. Paulin, S. Laurinavicius, A. Domanska, P. Somerharju and D. H. Bamford, “New, closely related haloarchaeal viral elements with different nucleic acid types”, *J. Virol.* **84** (2010), no. 7, pp. 3682–3689.
- Roitman, S., A. Rozenberg, T. Lavy, C. P. D. Brussaard, O. Kleifeld and O. Béjà, “Isolation and infection cycle of a polinton-like virus virophage in an abundant marine alga”, *Nat. Microbiol.* **8** (2023), no. 2, pp. 332–346.
- Roossinck, M. J. and E. R. Bazan, “Symbiosis: viruses as intimate partners”, *Annu. Rev. Virol.* **4** (2017), no. 1, pp. 123–139.
- Roux, S. and C. Coclet, “Viromics approaches for the study of viral diversity and ecology in microbiomes”, *Nat. Rev. Genet.* **27** (2026), no. 1, pp. 32–46.
- Roux, S., F. Enault, G. Bronner, D. Vaultot, P. Forterre and M. Krupovic, “Chimeric viruses blur the borders between the major groups of eukaryotic single-stranded DNA viruses”, *Nat. Commun.* **4** (2013), article no. 2700.
- Roux, S., M. G. Fischer, T. Hackl, L. A. Katz, F. Schulz and N. Yutin, “Updated virophage taxonomy and distinction from Polinton-like viruses”, *Biomolecules* **13** (2023), no. 2, article no. 204.
- Roux, S., M. Krupovic, R. A. Daly, et al., “Cryptic inoviruses revealed as pervasive in bacteria and archaea across Earth’s biomes”, *Nat. Microbiol.* **4** (2019), no. 11, pp. 1895–1906.
- Roux, S., M. Krupovic, A. Poulet, D. Debroas and F. Enault, “Evolution and diversity of the Microviridae viral family through a collection of 81 new complete genomes assembled from virome reads”, *PLoS One* **7** (2012), no. 7, article no. e40418.
- Sahakyan, H., K. S. Makarova and E. V. Koonin, “Search for origins of anti-CRISPR proteins by structure comparison”, *CRISPR J.* **6** (2023), no. 3, pp. 222–231.
- Santos-Perez, I., D. Charro, D. Gil-Carton, M. Azkargorta, F. Elortza, D. H. Bamford, H. M. Oksanen and N. G. A. Abrescia, “Structural basis for assembly of vertical single beta-barrel viruses”, *Nat. Commun.* **10** (2019), no. 1, article no. 1184.
- Sauguet, L., “The extended “two-barrel” polymerases superfamily: structure, function and evolution”, *J. Mol. Biol.* **431** (2019), no. 20, pp. 4167–4183.
- Sayers, E. W., J. Beck, E. E. Bolton, et al., “Database resources of the National Center for Biotechnology Information in 2025”, *Nucleic Acids Res.* **53** (2025), no. D1, pp. D20–D29.
- Seed, K. D., D. W. Lazinski, S. B. Calderwood and A. Camilli, “A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity”, *Nature* **494** (2013), no. 7438, pp. 489–491.
- Segel, M., B. Lash, J. Song, et al., “Mammalian retrovirus-like protein PEG10 packages its own mRNA and can be pseudotyped for mRNA delivery”, *Science* **373** (2021), no. 6557, pp. 882–889.
- Sencilo, A., D. Jacobs-Sera, D. A. Russell, et al., “Snapshot of haloarchaeal tailed virus genomes”, *RNA Biol.* **10** (2013), no. 5, pp. 803–816.
- Senkevich, T. G., N. Yutin, Y. I. Wolf, E. V. Koonin and B. Moss, “Ancient gene capture and recent gene loss shape the evolution of orthopoxvirus–host interaction genes”, *mBio* **12** (2021), no. 4, article no. e0149521.
- Shannon-Lowe, C. and A. Rickinson, “The global landscape of EBV-associated tumors”, *Front. Oncol.* **9** (2019), article no. 713.
- Shao, Q., I. V. Agarkova, E. A. Noel, et al., “Near-atomic, non-icosahedrally averaged structure of giant virus Paramecium bursaria chlorella virus 1”, *Nat. Commun.* **13** (2022), no. 1, article no. 6476.
- Sharon, I., A. Alperovitch1, F. Rohwer, et al., “Photosystem I gene cassettes are present in marine virus genomes”, *Nature* **461** (2009), no. 7261, pp. 258–262.
- Al-Shayeb, B., R. Sachdeva, L. X. Chen, et al., “Clades of huge phages from across Earth’s ecosystems”, *Nature* **578** (2020), no. 7795, pp. 425–431.
- Al-Shayeb, B., P. Skopintsev, K. M. Soczek, et al., “Diverse virus-encoded CRISPR-Cas systems include streamlined genome editors”, *Cell* **185** (2022), no. 24, pp. 4574–4586.
- Shepherd, J. D., “Arc - an endogenous neuronal retrovirus?”, *Semin. Cell Dev. Biol.* **77** (2018), pp. 73–78.
- Shkorporov, A. N. and C. Hill, “Bacteriophages of the human gut: the “known unknown” of the microbiome”, *Cell Host Microb.* **25** (2019), no. 2, pp. 195–209.

- Simmonds, P., M. J. Adams, M. Benko, et al., “Consensus statement: virus taxonomy in the age of metagenomics”, *Nat. Rev. Microbiol.* **15** (2017), no. 3, pp. 161–168.
- Simmonds, P., E. M. Adriaenssens, E. J. Lefkowitz, et al., “Changes to virus taxonomy, the international code of virus classification and nomenclature, and the ICTV statutes ratified by the International Committee on Taxonomy of Viruses (2025)”, *Arch. Virol.* **171** (2025), no. 1, article no. 23.
- Simmonds, P., E. M. Adriaenssens, F. M. Zerbini, et al., “Four principles to establish a universal virus taxonomy”, *PLoS Biol.* **21** (2023), no. 2, article no. e3001922.
- Solden, L., K. Lloyd and K. Wrighton, “The bright side of microbial dark matter: lessons learned from the uncultivated majority”, *Curr. Opin. Microbiol.* **31** (2016), pp. 217–226.
- Souza, F., R. Rodrigues, E. Reis, M. Lima, B. La Scola and J. Abrahao, “In-depth analysis of the replication cycle of Orpheovirus”, *Viol. J.* **16** (2019), no. 1, article no. 158.
- Strömsten, N. J., D. H. Bamford and J. K. Bamford, “In vitro DNA packaging of PRD1: a common mechanism for internal-membrane viruses”, *J. Mol. Biol.* **348** (2005), no. 3, pp. 617–629.
- Szathmary, E., “Toward major evolutionary transitions theory 2.0”, *Proc. Natl. Acad. Sci. USA* **112** (2015), no. 33, pp. 10104–10111.
- Szathmary, E. and J. Maynard Smith, “From replicators to reproducers: the first major transitions leading to life”, *J. Theor. Biol.* **187** (1997), no. 4, pp. 555–571.
- Szirovicza, L., U. Hetzel, A. Kipar, L. Martinez-Sobrido, O. Vapalahti and J. Hepojoki, “Snake deltavirus utilizes envelope proteins of different viruses to generate infectious particles”, *mBio* **11** (2020), no. 2, article no. e03250-19.
- Tamarit, D., E. F. Caceres, M. Krupovic, R. Nijland, L. Eme, N. P. Robinson and T. J. G. Ettema, “A closed Candidatus Odinarcahium chromosome exposes Asgard archaeal viruses”, *Nat. Microbiol.* **7** (2022), no. 7, pp. 948–952.
- Telengech, P., K. Hyodo, H. Ichikawa, R. Kuwata, H. Kondo and N. Suzuki, “Replication of single viruses across the kingdoms, fungi, plantae, and animalia”, *Proc. Natl. Acad. Sci. USA* **121** (2024), no. 25, article no. e2318150121.
- Tobiasson, V., J. Luo, Y. I. Wolf and E. V. Koonin, “Dominant contribution of Asgard archaea to eukaryogenesis”, *Nature* **650** (2026), no. 8100, pp. 141–149.
- Toro, N. and R. Nisa-Martinez, “Comprehensive phylogenetic analysis of bacterial reverse transcriptases”, *PLoS One* **9** (2014), no. 11, article no. e114083.
- Turner, D., A. N. Shkoporov, C. Lood, et al., “Abolishment of morphology-based taxa and change to binomial species names: 2022 taxonomy update of the ICTV bacterial viruses subcommittee”, *Arch. Virol.* **168** (2023), no. 2, article no. 74.
- Urayama, S. I., A. Fukudome, M. Hirai, et al., “Double-stranded RNA sequencing reveals distinct riboviruses associated with thermoacidophilic bacteria from hot springs in Japan”, *Nat. Microbiol.* **9** (2024), no. 2, pp. 514–523.
- Urayama, S. I., A. Fukudome, P. Mutz, et al., “Identification of hot spring Obelisk-like RNA replicons and expanded diversity of the Obelisk superfamily”, *Nat. Commun.* **17** (2026), no. 1, article no. 3041.
- van den Berg, D. F. and S. J. J. Brouns, “Phage tRNAs: decoding the enigma”, *Trends Microbiol.* **33** (2025), no. 10, pp. 1121–1131.
- van Oers, M. M., E. A. Herniou, J. A. Jehle, et al., “Developments in the classification and nomenclature of arthropod-infecting large DNA viruses that contain pif genes”, *Arch. Virol.* **168** (2023), no. 7, article no. 182.
- Vardi, A., B. A. Van Mooy, H. F. Fredricks, K. J. Popendorf, J. E. Osolinski, L. Haramaty and K. D. Bidle, “Viral glycosphingolipids induce lytic infection and cell death in marine phytoplankton”, *Science* **326** (2009), no. 5954, pp. 861–865.
- Varsani, A., B. Harrach, P. Roumagnac, et al., “2024 taxonomy update for the family *Circoviridae*”, *Arch. Virol.* **169** (2024), no. 9, article no. 176.
- Walker, P. J., S. G. Siddell, E. J. Lefkowitz, et al., “Recent changes to virus taxonomy ratified by the International Committee on Taxonomy of Viruses (2022)”, *Arch. Virol.* **167** (2022), no. 11, pp. 2429–2440.
- Wang, D., “The enigma of picobirnaviruses: viruses of animals, fungi, or bacteria?”, *Curr. Opin. Virol.* **54** (2022), article no. 101232.
- Wang, F., D. P. Baquero, L. C. Beltran, et al., “Structures of filamentous viruses infecting hyperthermophilic archaea explain DNA stabilization in extreme environments”, *Proc. Natl. Acad. Sci. USA* **117** (2020), no. 33, pp. 19643–19652.
- Wang, F., D. P. Baquero, Z. Su, T. Osinski, D. Prangishvili, E. H. Egelman and M. Krupovic, “Structure of a filamentous virus uncovers familial ties within the archaeal virosphere”, *Virus Evol.* **6** (2020), no. 1, article no. veaa023.
- Wang, F., V. Cvirkaite-Krupovic, M. Vos, et al., “Spindle-shaped archaeal viruses evolved from rod-shaped ancestors to package a larger genome”, *Cell* **185** (2022), no. 8, pp. 1297–1307.
- Wang, F., Y. Liu, Z. Su, et al., “A packing for A-form DNA in an icosahedral virus”, *Proc. Natl. Acad. Sci. USA* **116** (2019), no. 45, pp. 22591–22597.
- Wang, H. C., C. C. Chou, K. C. Hsu, C. H. Lee and A. H. Wang, “New paradigm of functional regulation by DNA mimic proteins: recent updates”, *IUBMB Life* **71** (2019), no. 5, pp. 539–548.
- Wang, Y., Z. Duan, Z. Chen, et al., “A new family of rolling circle replication endonucleases widespread in archaeal viruses and plasmids”, *Nucleic Acids Res.* **54** (2026), no. 8, article no. gkag363.
- Weigel, C. and H. Seitz, “Bacteriophage replication modules”, *FEMS Microbiol. Rev.* **30** (2006), no. 3, pp. 321–381.
- Wigington, C. H., D. Sonderegger, C. P. Brussaard, et al., “Re-examination of the relationship between marine virus and microbial cell abundances”, *Nat. Microbiol.* **1** (2016), article no. 15024.
- Wille, M., H. J. Netter, M. Littlejohn, et al., “A divergent hepatitis D-like agent in birds”, *Viruses* **10** (2018), no. 12, article no. 720.
- Williams, T. A., T. M. Embley and E. Heinz, “Informational gene phylogenies do not support a fourth domain of life for nucleocytoplasmic large DNA viruses”, *PLoS One* **6** (2011), no. 6, article no. e21080.
- Wolf, Y. I., D. Kazlauskas, J. Iranzo, A. Lucia-Sanz, J. H. Kuhn, M. Krupovic, V. V. Dolja and E. V. Koonin, “Origins and evolution of the global RNA virome”, *MBio* **9** (2018), no. 6, article no. e02329-18.
- Wolf, Y. I., S. Silas, Y. Wang, et al., “Doubling of the known set of RNA viruses by metagenomic analysis of an aquatic virome”, *Nat. Microbiol.* **5** (2020), no. 10, pp. 1262–1270.
- Wynne, S. A., R. A. Crowther and A. G. Leslie, “The crystal structure of the human hepatitis B virus capsid”, *Mol. Cell* **3** (1999), no. 6, pp. 771–780.

- Xiong, Y. and T. H. Eickbush, "Origin and evolution of retroelements based upon their reverse transcriptase sequences", *EMBO J.* **9** (1990), no. 10, pp. 3353–3362.
- Yan, Y., J. Zheng, X. Zhang and Y. Yin, "dbAPIS: a database of anti-prokaryotic immune system genes", *Nucleic Acids Res.* **52** (2023), no. D1, pp. D419–D425.
- Yang, Y., L. Delcourte, C. van Belleghem, et al., "Structure and nucleic acid interactions of the S(Delta60) domain of the hepatitis delta virus small antigen", *Proc. Natl. Acad. Sci. USA* **122** (2025), no. 19, article no. e2411890122.
- Yoshida, M., S. Medvedeva, A. Fukudome, et al., "'Paraxenoviridae', a putative family of globally distributed marine bacteriophages with double-stranded RNA genomes", *ISME J.* **19** (2025), no. 1, article no. wraf139.
- Yutin, N., P. Mutz, M. Krupovic and E. V. Koonin, "Mriyaviruses: small relatives of giant viruses", *mBio* **15** (2024), no. 7, article no. e0103524.
- Yutin, N., I. Tolstoy, P. Mutz, Y. I. Wolf, M. Krupovic and E. V. Koonin, "DNA polymerase swapping in *Caudoviricetes* bacteriophages", *Viol. J.* **21** (2024), no. 1, article no. 200.
- Yutin, N., Y. I. Wolf and E. V. Koonin, "Origin of giant viruses from smaller DNA viruses not from a fourth domain of cellular life", *Virology* **466–467** (2014), pp. 38–52.
- Zayed, A. A., J. M. Wainaina, G. Dominguez-Huerta, et al., "Cryptic and abundant marine viruses at the evolutionary origins of Earth's RNA virome", *Science* **376** (2022), no. 6589, pp. 156–162.
- Zhang, R., L. Mayer, H. Hikida, Y. Shichino, M. Mito, A. Willemsen, S. Iwasaki and H. Ogata, "A giant virus forms a specialized subcellular environment within its amoeba host for efficient translation", *Nat. Microbiol.* **11** (2026), no. 2, pp. 584–596.
- Zhang, S., M. Chu and X. Sun, "The arms race in bacteriophage interaction: deciphering bacteria defense and phage anti-defense mechanisms through metagenomics", *Front. Microbiol.* **16** (2025), article no. 1687307.
- Zhang, Y., Y. Guo, T. Qiu, M. Gao and X. Wang, "Bacteriophages: Underestimated vehicles of antibiotic resistance genes in the soil", *Front. Microbiol.* **13** (2022), article no. 936267.
- Zhao, Y., Y. Ma, C. Vasileiou, A. D. Farr, D. W. Rogers and P. B. Rainey, "Jumbo phage-mediated transduction of genomic islands", *Proc. Natl. Acad. Sci. USA* **122** (2025), no. 44, article no. e2512465122.
- Zheludev, I. N., R. C. Edgar, M. J. Lopez-Galiano, M. de la Pena, A. Babaian, A. S. Bhatt and A. Z. Fire, "Viroid-like colonists of human microbiomes", *Cell* **187** (2024), no. 23, pp. 6521–6536.
- Zhong, K. X., A. M. Chan, K. M. Miller, R. Saunders and C. A. Suttle, "Evolutionarily divergent nidovirus with an exceptionally large genome identified in Pacific oysters undergoing mass mortality", *Proc. Natl. Acad. Sci. USA* **122** (2025), no. 33, article no. e2426923122.
- Zhou, Y., A. Gutierrez-Preciado, Y. Liu, D. Moreira, M. M. Yakimov, P. Lopez-Garcia and M. Krupovic, "Viruses and virus satellites of haloarchaea and their nanosized DPANN symbionts reveal intricate nested interactions", *Nat. Microbiol.* **10** (2025), no. 11, pp. 2673–2685.
- zur Hausen, H., "Papillomaviruses and cancer: from basic studies to clinical application", *Nat. Rev. Cancer* **2** (2002), no. 5, pp. 342–350.