



## OPEN ACCESS

## EDITED BY

Ana Grande-Pérez,  
Instituto de Hortofruticultura  
Subtropical y Mediterránea "La Mayora"  
(IHSM-UMA-CSIC), Spain

## REVIEWED BY

Manasvi Lingam,  
Florida Institute of Technology,  
United States  
Richard Allen White,  
University of North Carolina  
at Charlotte, United States

## \*CORRESPONDENCE

Ester Lázaro  
lazarole@cab.inta-csic.es

## SPECIALTY SECTION

This article was submitted to  
Virology,  
a section of the journal  
Frontiers in Microbiology

RECEIVED 31 August 2022

ACCEPTED 10 October 2022

PUBLISHED 26 October 2022

## CITATION

de la Higuera I and Lázaro E (2022)  
Viruses in astrobiology.  
*Front. Microbiol.* 13:1032918.  
doi: 10.3389/fmicb.2022.1032918

## COPYRIGHT

© 2022 de la Higuera and Lázaro. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Viruses in astrobiology

Ignacio de la Higuera<sup>1</sup> and Ester Lázaro<sup>2\*</sup>

<sup>1</sup>Department of Biology, Center for Life in Extreme Environments, Portland State University, Portland, OR, United States, <sup>2</sup>Centro de Astrobiología (CAB), CSIC-INTA, Torrejón de Ardoz, Spain

Viruses are the most abundant biological entities on Earth, and yet, they have not received enough consideration in astrobiology. Viruses are also extraordinarily diverse, which is evident in the types of relationships they establish with their host, their strategies to store and replicate their genetic information and the enormous diversity of genes they contain. A viral population, especially if it corresponds to a virus with an RNA genome, can contain an array of sequence variants that greatly exceeds what is present in most cell populations. The fact that viruses always need cellular resources to multiply means that they establish very close interactions with cells. Although in the short term these relationships may appear to be negative for life, it is evident that they can be beneficial in the long term. Viruses are one of the most powerful selective pressures that exist, accelerating the evolution of defense mechanisms in the cellular world. They can also exchange genetic material with the host during the infection process, providing organisms with capacities that favor the colonization of new ecological niches or confer an advantage over competitors, just to cite a few examples. In addition, viruses have a relevant participation in the biogeochemical cycles of our planet, contributing to the recycling of the matter necessary for the maintenance of life. Therefore, although viruses have traditionally been excluded from the tree of life, the structure of this tree is largely the result of the interactions that have been established throughout the intertwined history of the cellular and the viral worlds. We do not know how other possible biospheres outside our planet could be, but it is clear that viruses play an essential role in the terrestrial one. Therefore, they must be taken into account both to improve our understanding of life that we know, and to understand other possible lives that might exist in the cosmos.

## KEYWORDS

astrobiology, virosphere, origin of life, quasispecies, horizontal gene transfer, virus biosignatures, experimental evolution

## Introduction

Astrobiology is a scientific discipline concerned with the origin, evolution, distribution and future of life in the universe (Des Marais et al., 2008; Hays et al., 2015; Cockell, 2020; O'rourke et al., 2020; Lingam and Loeb, 2021; Bennett et al., 2022). The use of the term life in reference to the object of study of astrobiology may raise doubts

about whether viruses should be included in its field of research. Some of the difficulties in deciding this question derive from the absence of a universal definition of life (Dix, 2003; Benner, 2010; Cleland, 2011, 2019), which is partly due to the fact that the only example of life we know is that on Earth. That all living beings on this planet share the same ancestor (Woese et al., 1990) makes it difficult to identify the fundamental properties of life. What we think is essential because of its presence in all forms of known life, may simply represent a characteristic that has been inherited from the common progenitor.

NASA defines life as a self-sustaining chemical system capable of darwinian evolution. Although viruses are chemical systems composed of the same molecules as life, they cannot use matter and energy from the external environment to build internal order in the same way that life does. Thus, if we regard the concept of self-sustaining system, viruses should not be considered living entities (Moreira and López-García, 2009). Nevertheless, the fact that viruses contain a genome that encodes proteins following the same rules as cells –together with the reality that they can reproduce and evolve– are sufficient arguments for other scientists to consider them as life (Hegde et al., 2009; Navas-Castillo, 2009; Forterre, 2016; Harris and Hill, 2020). To further complicate the question, some recently discovered viruses have genomes and dimensions comparable to some cellular microorganisms (Van Etten and Meints, 1999; La Scola et al., 2003; Abergel and Claverie, 2020) and can even be infected by other viral entities called virophages (Fischer, 2011; Mougari et al., 2019).

If we go back to the times when life was taking its first steps, the boundaries between life and non-life were more blurred than they are today (Szostak, 2012). Although most life definitions contain both a thermodynamic and an inheritance aspect, it is easy to imagine that at primitive times there were entities in which these two properties were not present in the same way they are in modern life. An example of these gray zones would be represented by autocatalytic networks, in which all molecules can be synthesized through the reactions catalyzed within the set (Vitas and Dobovišek, 2019), giving rise to a self-maintained system. In these networks, information is not stored in homopolymers such as DNA or RNA and, therefore, evolution would reside in variations in inter-molecular interactions or in the mechanisms of obtaining energy from the environment. Viruses would be another example of gray zone that can exist as an inert form, the virion, and an active form, the virus multiplying inside the cell. In their active form, viruses are self-organized systems that store and transmit information, and are able to maintain their organization despite changes in the environment. Although viruses always need to infect a cell to reproduce, all the information necessary to manipulate the cell metabolism and produce a viral progeny is contained in the virus genome itself, which brings them closer to the definition of life. Forterre introduced the concept of virocell (Forterre, 2011) as a living

system corresponding to any cell infected by a virus, whose main function is to produce new virions that would act as seeds or spores.

Whether viruses are considered living things or not, there is no doubt that finding viruses on a planet other than Earth would immediately make us think about the possibility of the existence of life on it. That life could be well established, taking its first steps, or even be extinct, having left viruses as the last vestiges of its existence. Viruses are inseparable companions of life as we know it. As we will discuss extensively in this review, viruses are probably very ancient and were present in the pool of genetic elements that gave rise to life (Koonin et al., 2006; Forterre and Prangishvili, 2009; Villarreal and Witzany, 2010; Durzyńska and Goździcka-Józefiak, 2015; Krupovic et al., 2019). Since that time, they have influenced, and continue influencing, evolution and distribution of life (Filée et al., 2003; Forterre, 2006; Koonin et al., 2006; Koonin, 2016). And there is no doubt of their relevance to conform the properties of the life that will populate our planet in the future.

Traditionally, viruses have been seen as disease-causing agents, a view that has changed radically in recent years. Pathogenic viruses are only a small fraction of the wide variety of viruses that exist. Many of them coexist peacefully with their hosts, sometimes even causing benefits. Often, advantages provided by viruses derive from their high capacity to modify and exchange pieces of their genomes, which convert them in great inventors of genes that, thanks to horizontal gene transfer, can be subsequently transferred to the cellular world (Muniesa et al., 2011; Penadés et al., 2015; Gilbert and Cordaux, 2017; Weiss, 2017; Irwin et al., 2022), giving rise to many of the intricate connections among the different branches of the phylogenetic tree of life.

Viruses exist practically everywhere in our planet, including extreme environments (Sawström et al., 2008; Yoshida-Takashima et al., 2012; Munson-McGee et al., 2020), which in some cases have physicochemical conditions that resemble some of those present in extraterrestrial environments. They are the most abundant biological entities on Earth and it is believed that all cellular organisms can be infected by some kind of virus (Suttle, 2005; Mushegian, 2020), which raises the question of whether they are an inevitable consequence of the emergence of living systems (Iranzo et al., 2016; Koonin et al., 2017) and must be present in any potential biosphere. Viruses, in the virion form, can withstand more extreme conditions than most cellular life. For this reason, sometimes they have even been considered as possible containers for the transport of genetic material between planets (Griffin, 2013; Berliner et al., 2018).

Finally, the rapidity of viral evolution is being exploited in multiple laboratories to carry out studies devoted to the research of biological evolution in real time (Elena et al., 2008; Kawecki et al., 2012; McDonald, 2019), which allows to establish some of the general principles governing this process. The importance

in evolution of the error rate value, the intensity of selective pressures, the population size or the standing genetic diversity are just some examples of the topics that can be studied using viruses as experimental system.

The conclusion that arises from all the considerations described above is that, although viruses are not considered *sensu stricto* life, both current and past life in our planet has been shaped thanks to their action. As knowledge has advanced, viruses have gone from being considered enemies of life to being recognized highly relevant in fields as diverse as molecular ecology, evolutionary biology, structural biology or geomicrobiology. The same is happening with astrobiology, which is increasingly realizing that to understand life in a wide sense it is absolutely necessary to include viruses in its scenario (Griffin, 2013; Berliner et al., 2018).

## What can viruses teach us about the origin of life?

How old are viruses? Does their simplicity mean that they are intermediate forms between inert and living matter? Or do they represent an alternative path followed by the same precursors that gave rise to life? Some proposals also state that viruses are modern entities that have derived from cells. So far, we do not have a definite answer for these questions, although most of the evidence indicates that viruses, particularly those with an RNA genome, are very old. Therefore, their study can provide relevant information about the first steps toward life.

## Origin of viruses

There are three main hypotheses to explain the origin of viruses (Forterre, 2006; Koonin et al., 2006; Forterre and Prangishvili, 2009; Krupovic et al., 2019) that are briefly described below. Each of them has its implications in the relationships that the viral world establishes with the cellular world and in their mutual influence throughout life history. The three hypotheses are not exclusive in the sense that different groups of viruses could have been originated through different routes and at different times. For example, viruses with a DNA genome could not have existed until the emergence of this molecule, which places their origin at a later time than that of RNA viruses.

### Escape hypothesis

Viruses are genetic cell elements that escaped from cells, becoming infectious entities unable to replicate on their own. According to this idea, prior to the emergence of DNA cells, there was a stage in the evolution of life that was dominated by RNA cells containing fragmented genomes (Woese, 1987, 2002). In contrast to the current world –in which the processes

of genomic replication and cell division are perfectly integrated– in the primitive world, there were probably no mechanisms that regulated this integration in a precise way. Thus, it would be easier at that time that a genomic fragment could be released from a cell and become an infectious unit capable of replicating at the expense of the resources produced by others. For this hypothesis to be true, there should be more similarities between the genes of viruses that infect a particular domain of life and the cellular genes of that same domain. However, when virus protein structural domains have been analyzed, in many cases they are found to be similar in bacteriophages, archaeal viruses and eukaryotic viruses (Bamford, 2003; Benson et al., 2004; Rice et al., 2004; Trus et al., 2004; Krupovic and Bamford, 2008; Abrescia et al., 2012; Brum et al., 2013).

### Reductive hypothesis

According to this scenario, viruses are the product of the degeneration of ancestral cells that lost their machinery for protein synthesis and energy production. These deteriorated cells became parasites of others that kept all their capacities. No evolutionary intermediary between viruses and cells has been found to date. Moreover, when viruses are compared with other parasites, the latter always retain some characteristics of their free-living time. The discovery of giant viruses that infect protists, and that sometimes possess genes of the protein translation machinery, triggered a resurgence of the reductive hypothesis (Schulz et al., 2017; Abrahão et al., 2018; Rodrigues et al., 2020). However, it is now widely accepted that these genes have been acquired from the host (Koonin and Yutin, 2018; Brahim Belhaouari et al., 2022).

### Virus-first hypothesis

This hypothesis states that viruses, particularly those with RNA genomes, descend from the first molecules with replicative capacity that likely existed on Earth before the appearance of cellular life. For a time, this thinking received little consideration, because actual viruses always need a cell to replicate and, therefore, cells should have preceded viruses. However, primitive viruses were probably very different from today's viruses. They could have been mere genetic parasites that emerged in replicator networks endowed with catalytic capacities. Parasites would multiply at the expense of the products of the network without contributing to their generation, as it is shown in some theoretical studies (Iranzo et al., 2016; Koonin et al., 2017). Additional, indirect support for this hypothesis comes from the fact that viruses use a great diversity of molecules [single-stranded RNA (ssRNA), double-stranded RNA (dsRNA), single-stranded DNA (ssDNA), and double-stranded DNA (dsDNA)] and strategies to store and replicate their genetic information, all of which brings to mind a time when different ways of preserving and processing genetic information were being tested (Koonin et al., 2006; Holmes, 2011b; Krupovic et al., 2019).

## The RNA world

How was the pre-cellular state of biological evolution where viruses could first emerge? In modern life, functional proteins can be synthesized because DNA stores information about the order in which amino acids should be linked together. But going from the DNA sequence to the sequence of a protein is an intricate process, which in turn requires the intervention of other proteins and complex structures such as ribosomes. Therefore, separation of information and function into two different molecules poses a paradox that can only be resolved if, in primitive life, information and function resided in the same molecule. Many scientists accept that at some stage of the evolution of life this molecule was RNA, although this does not preclude the existence of pre-RNA worlds in which catalysis and information could reside in other molecules that were more stable or easier to assemble than RNA. In this context, PNAs (peptide nucleic acids), which are synthetic polymers with a simple, achiral chemical structure composed of repeating N-(2-aminoethyl)-glycine units linked by peptide bonds (Nielsen, 2007) have received much consideration.

That RNA can be used to store hereditary information is demonstrated by the existence of viroids and RNA viruses. Due to internal base pairing, RNA molecules fold into secondary and tertiary structures that maximize their stability. As it happens with proteins in the current world, these structures can lead to the formation of active centers that facilitate the catalysis of certain chemical reactions. RNA molecules with catalytic capacity are called ribozymes. Their existence was first demonstrated in the eighties (Kruger et al., 1982; Bass and Cech, 1984; Altman, 1989) and confirmed in many subsequent studies (Chapman and Szostak, 1994; Hager et al., 1996; Lee and Lee, 2017; Janzen et al., 2020).

Everything described above led to postulate the existence of a hypothetical RNA world, made up of ensembles of molecules, the so-called primitive replicators, capable of storing and transmitting information (Gilbert, 1986; Oro et al., 1990; Sankaran, 2016). The catalytic properties of these molecules could have facilitated the emergence of a simple metabolism, which, once individualized in a compartment, would have been the basis for the appearance of the first cells (Koonin et al., 2006; Koonin, 2014; Szostak, 2017; Joyce and Szostak, 2018) and probably of the first parasites.

The first experiment demonstrating the capacity of RNA molecules to evolve in response to the environment was carried out by Sol Spiegelman in 1967 (Mills et al., 1967), using the bacteriophage Q $\beta$ , a virus with an RNA genome (Figure 1). The experiment carried out by Spiegelman consisted in mixing in a test tube a small amount of viral RNA with all the components necessary for its replication. This mixture was incubated under optimal conditions for the time necessary for the RNA to be copied. Then, a fraction of the molecules produced were transferred to a new tube containing fresh

substrates for replication. This process of serial transfers was repeated over time, with the aim of finding out what happened to the RNA molecules. The result was that the virus genome lost a large part of its sequence, which allowed it to increase the copying rate (Figure 1). In new experiments it was found that, if the environment changed, the RNA molecules were capable of specific adaptations, increasing their replication speed under conditions such as imbalanced nucleotide concentrations or presence of ribonucleases (Levisohn and Spiegelman, 1969; Kacian et al., 1972; Biebricher and Gardiner, 1997). These experiments were highly relevant to demonstrate that molecules are also susceptible of optimization through darwinian evolution, a concept of crucial importance for the idea that before cellular life there must have been a phase of molecular evolution.

Theoretical studies inspired in Spiegelman's experiments showed that error-prone replication combined with natural selection, and acting for a sufficiently long time in an infinite population of replicators (RNA-like molecules) differing in their fitness values, would result in the generation of a steady-state distribution of mutants in which each of them has a constant frequency (Eigen, 1971; Eigen and Schuster, 1979). This population structure was referred to as quasispecies, and the framework describing its dynamics quasispecies theory (Biebricher and Eigen, 2006). The major achievement of quasispecies theory was to describe mathematically a system involving multiplication of molecules with a regular production of error copies. The sequence displaying the highest replication rate within the quasispecies was denoted master sequence and corresponded to that most represented in the whole set of mutants, also known as mutant spectrum. The ensemble of all possible variations that can be generated from a genetic sequence is known as the sequence space. For a replicator of length  $l$  the sequence space is  $4^l$ , a number that can be enormous, even for short replicators. Quasispecies population structure was an important component of the catalytic hypercycle, a primitive model of organization of self-replicating molecules connected in a cyclic and autocatalytic fashion (Eigen and Schuster, 1979).

In addition to defining in quantitative terms a system of replicators that are copied with high error rate, quasispecies theory established the value of the error rate (the so-called error threshold) that was compatible with the conservation of genetic information (Biebricher and Eigen, 2005). When the error threshold is exceeded, the superiority of the master sequence disappears, all possible sequences become equally probable and the genetic information is lost.

Almost in parallel, experimental studies carried out with bacteriophage Q $\beta$  (Batschelet et al., 1976; Domingo et al., 1978) showed that this virus replicated with very high error rate and showed a great heterogeneity in its populations. The typical error rates of RNA viruses are on the order of  $10^{-6}$  to  $10^{-4}$  errors per nucleotide copied (Holland et al., 1982; Duffy et al., 2008; Sanjuán et al., 2010; Domingo et al., 2021).

For comparison, error rate values (expressed with the same units as above) are around  $10^{-6}$  for ssDNA viruses and in the ranges between  $10^{-8}$  and  $10^{-7}$  for dsDNA viruses,  $10^{-10}$  and  $10^{-9}$  for bacteria, and  $10^{-11}$  and  $10^{-10}$  for eukaryotes [(Gago et al., 2009) and references therein]. These high error rate values, together with the usually large sizes of viral populations, lead to the generation of highly diverse populations whose behavior could mimic that of the ensembles of replicators present in the RNA world and described in quasispecies theory. Nevertheless, factors such as the mechanism of genome replication, the variability and intensity of the selective pressures or the frequency of population bottlenecks are also highly relevant to determine the structure of a virus population and its evolutionary dynamics. Consequently, the substitution rate per nucleotide and year in some eukaryotic ssDNA viruses is similar to that of ssRNA viruses with the same genome length (Duffy et al., 2008).

## Viruses in the RNA world

If viruses are considered as simple genetic parasites, it is easy to think that their predecessors could have been present in the ensemble of primitive replicators that made up the RNA or the pre-RNA world. Their emergence would require the generation of mutants capable of being copied without contributing to the generation of the resources necessary for the copying process. The occurrence of these mutants would be relatively frequent, given the low fidelity of primitive replication (Krakauer and Sasaki, 2002). However, since parasites are not capable of a self-copying process, their permanence would require that they remain stable in the environment for long periods of time. This means that resistance to damage caused by external physicochemical variables probably constituted a relevant trait subjected to the action of natural selection in the RNA world, which later could have led to the emergence of viral capsids. The current existence of defective virus mutants—which lack the ability to encode all the proteins needed to complete their infectious cycle, but can give rise to a progeny when these proteins are provided by other viruses—corroborates the ease of the emergence of this kind of parasites (Huang, 1973; Cole, 1975; Grande-Pérez et al., 2005).

A valid approach to investigate whether viruses could originate in the RNA world is to examine the genes that encode specific viral functions, such as genome replication. In this case, there are two activities that are only found in the viral world and, more specifically, in RNA viruses. These are the RNA-dependent RNA polymerase (RdRp), and the reverse transcriptase (RT). In the cellular world, the RdRp and the RT perform very specialized functions, but are not involved in genome replication. It is striking that both enzymes share with DNA-dependent DNA polymerases a structural folding located in the main catalytic domain (Iyer et al., 2005; Kazlauskas et al., 2016), which seems

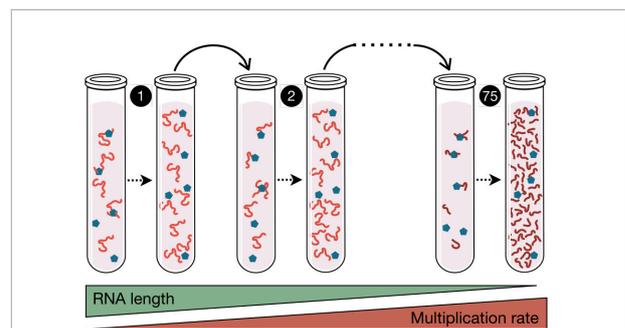


FIGURE 1

Spigelman's Monster. In this experiment published in 1967, the RNA from phage Q $\beta$  (red lines) and the enzyme responsible for its copy (blue pentagons) were incubated in the presence of nucleotides, and subjected to *in vitro* replication. The products of each polymerization reaction were serially transferred to new tubes. After 75 of these transfers, 83% of the RNA genome was eliminated, which in turn increased the multiplication rate of the RNA population. This seminal paper by Mills et al. (1967) demonstrated that darwinian evolution can occur in a test tube. Similar serial transfers experiments are used in the study of virus evolution.

to correspond to an ancestral RNA recognition motif (RRM) that was probably crucial for replication of both RNA and DNA (Cléry et al., 2008), serving as a cofactor for ribozymes in the RNA world. Other virus genes that also contain the RRM are the rolling-circle replication endonucleases, a superfamily of helicases, and some protein-primed DNA polymerases.

A relevant question is whether the primitive genetic parasites also encoded the proteins necessary for the formation of capsids. The current view is that virus capsid proteins evolved from cellular proteins at different stages of evolution (Krupovic et al., 2019). The high prevalence of some structural domains, such as the single jelly roll (a  $\beta$ -barrel fold consisting of eight antiparallel  $\beta$ -strands organized in two sheets that form the opposite sides of the barrel) or the double jelly roll in the capsids of viruses infecting organisms belonging to the three domains of life is probably due to their antiquity (Abrescia et al., 2012; Yutin et al., 2018; Santos-Pérez et al., 2019). In contrast to this, other capsid proteins may have been acquired more recently.

In conclusion, there are multiple lines of evidence indicating that the primitive RNA world could have given rise to two parallel and interconnected paths that led to the emergence of the first cells and the first viruses. Since then, both worlds have evolved in close link, influencing each other and contributing to increase their diversity, as we will see in the following sections.

## Role of viruses on the evolution of life

In the last decades, it has been recognized that viruses are major players in the history of life. They influence cellular evolution in different ways: acting as selective pressures,

behaving as vehicles for horizontal gene transfer, and creating new genes that provide evolutionary novelty.

## Viruses as selective pressure

Virus infections favor organisms that possess some type of defense mechanisms. In the prokaryotic world, these mechanisms include abortive infections, modification of the molecules used as receptors, or the selection of systems, such as restriction endonucleases or CRISPRs, that recognize the virus genetic material and degrade it (Gao et al., 2020; Hampton et al., 2020; Wang et al., 2020b). In turn, the appearance of microorganisms that better resist the infection exerts a selective pressure on the viral population that will favor viruses able to counteract the cellular defenses (Geoghegan and Holmes, 2018; Stanley and Maxwell, 2018; Koonin and Krupovic, 2020). The repetition of these cycles involves concerted changes in the pathogen and its host, an *arms race* that accelerates the evolution of both the virus and the host (Buckling and Brockhurst, 2012; Koskella and Brockhurst, 2014; Watson et al., 2021).

A similar scheme works for viruses infecting eukaryotic cells. In the particular case of vertebrates (Kaján et al., 2020), it has led to the selection of the adaptive immune system that, with a limited number of genes, is able to cope with a wide variety of antigens. It is believed that one of the possible causes of the emergence of this system has been the need to deal with highly changing agents such as viruses (Holmes, 2004). Sometimes viruses reduce their virulence in certain host species. In this way, they do not induce a strong immune response and can remain multiplying longer in the same host. When, on the one hand, a protective immune response to the virus is generated and, on the other hand, virulence is reduced, a pacific coexistence between the virus and its host can be reached (Barnard, 1984; Holmes and Duchêne, 2019). When one of these viruses that are well adapted to a particular animal species is able to infect the human population, a zoonosis occurs. The absence of immunity to the pathogen, together with the fact that the virulence has not been attenuated in the new host, can lead to the emergence of a new disease (French and Holmes, 2020). COVID-19, AIDS, or the pandemics caused by new variants of influenza virus are just some examples of diseases caused by zoonotic viruses. In the short term, the introduction of these or other viruses in the human population can be very negative, but in the long term they may contribute to make us stronger against new threats.

## Horizontal gene transfer in prokaryotes

Horizontal or lateral gene transfer allows innovations that arise in one group of organisms to be shared by a much larger set. Through transduction, both specialized and generalized, viruses are relevant agents of horizontal gene transfer in bacteria and archaea.

Specialized transduction takes place during the lysogenic cycle and implies that the virus, instead of destroying the cell, integrates its genome into the cell chromosome, being transferred to daughter cells by vertical transmission (Feiner et al., 2015). Most sequenced bacterial genomes contain at least one prophage (a phage genome integrated) and, in some cases, these elements constitute 10–20% of their total genetic material. Sometimes, after the lysogenic integration of a prophage, some signals occur that provoke its release, initiating a lytic cycle that will end with the production of new phages and the lysis of the bacterium. In the process, the phage frequently carries some cellular genes with it, which can be transferred to other bacteria in the course of new infections. Among the most frequently phage-transferred genes there are some that increase the virulence of the bacterium (Plunkett et al., 1999; Nishida et al., 2021), confer resistance to antibiotics (Gabashvili et al., 2020, 2022), help it to survive under difficult circumstances (Lindell et al., 2004, 2005; Sharon et al., 2007; Fridman et al., 2017), or allow the exploration of new ecological niches (Sullivan et al., 2005). In all cases, the advantages obtained from the integration of virus genetic material is what contributes to their persistence. Taking as an example the most abundant photosynthesizers on Earth –*Prochlorococcus* and *Synechococcus* (Flombaum et al., 2013)–, genes for the components of these cyanobacteria's photosystems are also present in the genomes of their phages (Mann et al., 2003), which seems to have amplified the evolution and expansion of these systems. Furthermore, the expression of the viral photosynthetic genes during infection enhances the photosynthetic activity of the host (Fridman et al., 2017), which increases viral replication, especially under low-nutrient conditions (Hurwitz et al., 2013).

Generalized transduction occurs during the lytic cycle of some viruses that, at the beginning of the infection, cause cellular DNA breakage (Thierauf et al., 2009; Waddell et al., 2009). When the progeny viruses are assembled, it is frequent that they do not distinguish between their own genome and the fragments of the cellular genome. The foreign DNA can be introduced into the virus capsids and, thanks to homologous recombination mechanisms, become integrated in the genome of the new infected microorganism.

## Virus endogenization

The insertion of virus genomes into the cell chromosome is not exclusive to the prokaryotic world. When it happens in the chromosomes of the host reproductive cells of vertebrates and the inserted provirus is maintained over time, the process is called viral endogenization (Holmes, 2011a; Feschotte and Gilbert, 2012; Greenwood et al., 2018). After the initial insertion event, the provirus can spread in the cellular genome through a mechanism similar to that used by retrotransposons. During their dissemination, the viral sequences accumulate

mutations or undergo epigenetic modifications that inactivate the expression of most of their genes (Wolf and Goff, 2008; Rowe and Trono, 2011; Yang et al., 2022), so that after a time they can no longer lead to the production of infectious viral particles.

The integrated viral elements can disrupt the sequence of genes or their regulatory elements, causing loss of function. Consequently, individuals carrying this kind of exogenous DNA are normally eliminated by natural selection. However, sometimes viral genomes persist through generations, expanding in the population until they become fixed. In these cases, the DNA of viral origin becomes host DNA. The great number of viral insertions that have been fixed in vertebrate lineages (Hayward et al., 2015) represents a substantial source of genetic material that adds variability to genomes and that sometimes can provide benefits. For instance, the insertion of a provirus in a region near the pancreatic amylase gene has allowed this enzyme to be also expressed in saliva in humans, so that starch digestion can begin in the mouth (Ting et al., 1992). This is just an example of how virus regulatory sequences can modify gene expression in the host. In addition to the effects on individual genes, the process of provirus amplification can lead to the dispersion of viral regulatory elements throughout the cellular genome, producing large reorganizations of gene expression that can produce major evolutionary innovations (Feschotte, 2008; Koonin, 2016). The amplification process also facilitates the exchange of genetic material between different sequences of the genome, resulting in the loss of some regions, the duplication of others, or changes in their location (Hughes and Coffin, 2001; Bai et al., 2021).

Despite the serious consequences that the expression of virus genes usually has for the host, there are some situations in which their products have been recruited to perform cellular functions. A classic example is the expression of viral proteins that provide immunity to infection by related viruses (Arnaud et al., 2008). But perhaps the best example of genes from endogenous viruses that have allowed the emergence of novel functions are those encoding for syncytins. Syncytins are proteins that promote the fusion of a type of cells called trophoblasts to give rise to the formation of the syncytiotrophoblast, a cell layer that forms part of the placenta. Syncytin genes have their origin in a gene from retroviruses, specifically in the gene *env*, which has fusogenic activity (Dupressoir et al., 2011). It appears that the event of domestication of *env* genes of viral origin has occurred at least six independent times throughout evolution, in different mammalian species, including humans (Rawn and Cross, 2008).

## Role of viruses in major evolutionary transitions

Darwinian theories explain evolution as the result of the progressive accumulation of heritable changes of small effect. However, these small changes cannot explain the increases

in biological complexity associated with some evolutionary innovations that require drastic changes in the amount of genetic information and/or in the way it is processed. When these innovations involve the integration of entities from a lower level of organization into a higher one that comes to constitute a new level of selection, we speak of major evolutionary transitions (Szathmáry and Smith, 1995; Szathmáry, 2015). The continuous arms-race between viruses and their hosts, together with the frequent exchange of genetic material driven by viruses, has frequently contributed to the increase of biological complexity (Forterre, 2006; Koonin, 2016; Mizuuchi et al., 2022), as illustrated in some of the examples given in the previous section.

It is believed that the need to defend against primitive viruses promoted that the molecules carrying genetic information became grouped in compartments that, in addition to create barriers for the spread of parasites, facilitated the cooperation between non-parasitic replicators, contributing to stabilize the whole system, which emerged as a new selection unit (Szathmáry and Demeter, 1987; Higgs and Lehman, 2015). Within compartments, selection for increased genome size would also have been generally favored, as a mechanism to fix the most favorable genetic combinations. Primitive compartments could be protective micro-environments in mineral surfaces or simple lipid vesicles.

Mathematical modeling of replicator systems also suggests that resistance to parasites increases when information and function are separated in different molecules. It is easy to understand the advantages of this fact, since parasites could only take advantage of the functional molecules and not of the informative ones, which would facilitate their survival. However, we still do not know much about how DNA replaced RNA as an information storage molecule and about the appearance of the first protein catalysts, which probably coexisted for some time with RNA or RNA-like catalysts. The presence in retroviruses of reverse transcriptase activity, which converts viral RNA to DNA, has led some authors (Forterre, 2006) to think that this enzyme could be involved in the replacement of RNA for DNA as informative molecule, but this is something that with the current evidence cannot be affirmed with certainty.

Viruses may also have played a role in the origin of eukaryotic cells, which possess mitochondria and an internal membrane system that delimits a nucleus in which the genomic DNA is located. Mitochondria have their origin in a process of endosymbiosis between a bacterium and another cell, probably an archaeon. It is intriguing that in all them, the bacterial RNA polymerase has been replaced by a phage RNA polymerase (Filée and Forterre, 2005), although it is difficult to elucidate the advantages provided by this fact. The eukaryotic nucleus allows the uncoupling between transcription and translation. The discovery of viral factories (Fridmann-Sirkis et al., 2016; Chaikerasitak et al., 2017), showing that some bacterial and eukaryotic viruses can establish a similar separation, led to

propose that the eukaryotic nucleus has a viral origin (Koonin, 2015; Bell, 2020; Takemura, 2020). However, this idea is not exempt of controversy and there are other hypotheses that state that the nucleus originated from invaginations of the plasmatic membrane in an ancestral prokaryote or has a symbiotic origin without virus participation.

## Viruses at the planetary scale: The virosphere

The ongoing advances in high-throughput sequencing technologies and metagenomic analyses have been crucial to reveal the immense diversity and pervasive distribution of viruses throughout our planet (Zhang et al., 2018). As previously indicated, viruses seem to have colonized – presumably associated with cellular hosts – every place where life can be sustained, while exploring vast areas of the evolutionary sequence space. The viral composition of the Earth's biosphere is known as the virosphere.

### The immensity of the virosphere and its diversity

The abundance of viruses in aquatic and terrestrial environments has been estimated to be 10–100-fold higher than that of unicellular organisms, although this ratio is dynamic and varies across ecological niches (Wigington et al., 2016). A milliliter of seawater can contain up to hundreds of millions of viruses (Bergh et al., 1989), and up to a billion of them can be found in a single gram of soil (Swanson et al., 2009). Thus, it has been estimated that the total number of viral particles in the planet is in the astronomical order of  $10^{31}$  (Hendrix et al., 1999; Mushegian, 2020), outnumbering all cellular organisms.

Metagenomic studies aimed at unveiling the viral makeup of different environments around the globe have barely scratched the surface of the virosphere's diversity. A remarkable effort is the deep sequencing of samples collected during the Tara Oceans expeditions, in which nearly 35,000 samples from the ocean were collected between 2009 and 2013 in a ship transect around the globe (Brum et al., 2015; Roux et al., 2016; Gregory et al., 2019; Sunagawa et al., 2020). This data set consisted mostly of viruses with dsDNA genomes (excluding large portions of the virosphere), and yet, they are taxonomically more abundant than bacteria, archaea and micro-eukaryotes together (Abreu et al., 2022).

The rate of discovery of RNA viruses has lagged behind that of DNA viruses, mostly due to the technical requirement of converting RNA to DNA prior to sequencing. However, thanks to advances in metatranscriptomics, there has been a recent explosion in the amount of RNA virus data. The collection of RNA viruses in the databases doubled in 2016 after the

metatranscriptomic analysis of numerous invertebrates (Shi et al., 2016), and doubled again in 2020 after sequencing just 10 L of sea water (Wolf et al., 2020). In 2022, by mining data from databases, the amount of known RNA virus sequences increased fivefold with the description of ~330,000 novel RNA viruses (Neri et al., 2022), in addition to the discovery of ~130,000 highly divergent RdRp sequences from novel viruses (Edgar et al., 2022). In each of these RNA virus data expansions, novel virus clades have been revealed. For instance, the exploration of the RNA virus oceanic composition has reached enough depth to reveal a novel group of capsid-less viruses, the *taraviricots*, which provides a missing link between retroelements and RNA viruses. The lineage of the *taraviricots* seems to predate the split between RdRPs and retrotranscriptases, suggesting that virus-like replicators encoding an RdRp are ancestral to RNA viruses and retroviruses, and probably derived from the primordial pool of replicators of the RNA world (Zayed et al., 2022).

### Organization of the virosphere and its challenges

Considering the colossal diversity of viruses, classifying the virosphere is not a trivial task. For decades, the main virus classification scheme was based on the nature of the packaged nucleic acid and traits such as virion morphology, replication strategy, host organism and type of disease. Given the explosion of virus sequence information fueled by metagenomics and for which no phenotypic information is available, these criteria seem no longer practical. The 7 Baltimore classes provided an organization framework based on the nature of the packaged nucleic acid and the strategies by which mRNAs get produced from these DNA or RNA genomes (Figure 2; Baltimore, 1971). However, the Baltimore classification does not accurately represent the evolutionary relationships among viral lineages.

A phylogenetic approach to virus classification is difficult, due to the absence of a common marker for virus classification. Moreover, viral genomes are modular and virus evolution is greatly shaped by lateral gene transfer. Despite these difficulties, by using phylogenetic analyses of virus hallmark genes and gene sharing networks, a new “megataxonomy” of viruses has recently been established (Koonin et al., 2020). In this new classification of viruses, a new taxonomic rank – the realm – had to be defined to distinguish between virus taxa with no common ancestry. The six virus realms are: (i) *Riboviria*, which contains all RNA viruses and retroviruses, and that may have derived from the primordial pool of replicons, (ii) *Monodnaviria*, with viruses whose DNA genomes (most single stranded) encode an endonuclease for replication, (iii) *Varidnaviria*, dsDNA viruses with vertical jelly roll capsid proteins, (iv) *Duplodnaviria*, dsDNA viruses with HK97-fold capsid proteins, (v) *Adnaviria*, which helical virions containing A-form DNA genomes, and (vi) *Ribozyviria*, whose members have circular RNA genomes



realm Monodnaviria appear to have emerged multiple times from plasmids by capturing capsid genes from RNA viruses (Kazlauskas et al., 2019).

## Impact of viruses on Earth's biogeochemical cycles and organismal biodiversity

Microbes are the main drivers of the biogeochemical cycles that define ecosystems, but they also are susceptible to virus infection, e.g., it is estimated that ~20% of the bacteria in the oceans are lysed by viruses per day (Suttle, 1994). These ongoing viral infections shape microbial community structures and are necessary for nutrient cycling (Dominguez-Huerta et al., 2022).

Viruses can impact food webs and biogeochemical cycles in different ways. Cell death caused by lytic viruses leads to the release of dissolved organic matter, which is then available for its consumption by low trophic levels. This process, called “viral shunt,” allows nutrients to be reused by other microorganisms rather than getting passed on to higher trophic levels (Wilhelm and Suttle, 1999; Zimmerman et al., 2020). The released organic matter from viral lysis can be particularly rich in nucleotides and amino acids (Ankrah et al., 2014). Thus, the viral shunt is important for the cycling of not just carbon, but other nutrients such as nitrogen (Shelford et al., 2012), phosphorus, and iron (Gobler et al., 1997; Poorvin et al., 2004). A complementary process, the “viral shuttle,” describes how viruses help the export of carbon by sinking organic matter to the bottom of the ocean (Guidi et al., 2016; Sullivan et al., 2017). The virus shuttle has been observed in algal blooms –often terminated through virus infection (Brussaard, 2004)–, in which the production of virally induced polymeric matrices enhances the sink of cellular aggregates to the sea floor (Laber et al., 2018). Virus composition have proven to be a great predictor of carbon flux in aquatic (Guidi et al., 2016; Kaneko et al., 2021) and terrestrial ecosystems (Emerson et al., 2018; Trubl et al., 2018; Starr et al., 2019), which further demonstrates the importance of virus infections in biogeochemical cycles.

The lysis of cells upon infection can have a significant effect on top-to-bottom control on biogeochemical cycles. However, viruses do not need to kill their host to impact ecosystems. During infection, viruses reprogram the cellular metabolism to divert resources to progeny production, which at times involves the expression of virally encoded metabolic genes that were likely acquired from cells (see section “Horizontal gene transfer in prokaryotes”) (Howard-Varona et al., 2020). Since a significant portion of the Earth's microbiome is infected at any given time, these viral “auxiliary metabolic genes” may carry functions that can have an impact on planetary processes (Zimmerman et al., 2020). The analysis of auxiliary metabolic genes from virus genomic data can be used as a proxy to understand the impact that viruses exert on ecosystems,

including the control of sulfur and nitrogen cycling in the ocean (Roux et al., 2016).

Viruses are also community structure regulators. They can propagate more effectively when their host is in abundance, thus, according to the “kill-the-winner” hypothesis, lytic viruses help maintain a balance and promote diversity in microbial communities (Thingstad and Lignell, 1997). This seems the case in the collapse of algal blooms by viral infection (Suttle, 2007). However, there are other examples where a particular species seems to be pervious to viral infection and become dominant (Yooseph et al., 2010). This scenario, called “king-of-the-mountain” (Giovannoni et al., 2013), suggests that high levels of recombination in these organisms allow the rapid adaptation of defense systems against their viruses (Zhao et al., 2013). This positive feedback loop in the co-evolution of virus and host is another example of the arms race described in section “Viruses as selective pressure,” by which viruses fuel organismal evolution and diversification (Giovannoni et al., 2013).

The realization of the important roles that viruses play in Earth's habitats underscore how pertinent it is to consider genetic parasites in the study of biological systems. For that matter, it is relevant to also look at those ecological niches where life defies its own physicochemical boundaries.

## Viruses in extreme environments

The study of life in harsh conditions is essential to astrobiology. On one hand, extremophilic organisms shed light on the limits of habitability, i.e., the range of environments capable of supporting life in terms of temperature, pH, salinity, pressure, humidity and radiation (An Astrobiology Strategy for the Search for Life in the Universe, 2019). On the other hand, the biology of extreme environments can help to understand the conditions for life emergence (Baaske et al., 2007; Morowitz and Smith, 2007; Weiss et al., 2016).

Extreme environments are no exception to the wide distribution of viruses. They can thrive at temperature ranges between –12 and 96°C (Munson-Mcgee et al., 2018), pH values from 1 to 11 (at the least), saturated salinity (Pietilä et al., 2013), and extreme dryness (Gil et al., 2021). According to this, viruses have been found in hydrothermal vents (Russell and Hall, 1997), including those located at great depths in the oceans (Thomas et al., 2021), in oil fields (Zheng et al., 2020), in the soils of hyper-arid deserts (Zablocki et al., 2015), at hundreds of meters in the subsurface (Daly et al., 2019), or in places with salt concentrations far above the optimal for most living beings (Santos et al., 2012). Many of these locations share environmental characteristics that also occur in extraterrestrial environments. For example, submarine hydrothermal vents could be considered a terrestrial analog of those postulated to exist at the bottom of the inner oceans of the icy moons of some gas giants, such as Enceladus, Europa, or Ganymede. Oil field

biology could help us to understand whether life could develop in the hydrocarbon lakes that exist on Titan, while the study of the subsurface could help us to figure out the possible life that existed (or still exists) on Mars. As we have already indicated, wherever there is cellular life there are also viruses, which reinforces the idea that the search for these entities could be used as a proxy for the search for life beyond Earth, something that is not receiving all the consideration it deserves.

Like in other ecosystems, extreme environment viruses can encode auxiliary metabolic genes that may be transferred between organisms. In hydrothermal vents, these genes are involved in metabolic pathways that are critical for survival in this environment including –but not limited to– sulfur-oxidizing enzymes (Anantharaman et al., 2014; Castelán-Sánchez et al., 2019). Viruses can also increase the adaptability of multicellular organisms to extreme environments. The triple-symbiosis in the thermotolerant terrestrial plant *Curvularia protuberata* is a classic example of how a virus infection –via fungal interaction in this case– can provide tolerance to heat stress (Márquez et al., 2007).

Archaea are typically found in the whole spectrum of extreme environments, and their viruses (archaeoviruses) –given the striking diversity of their unique morphologies (Figure 2)– are one of the most captivating types of extremophilic viruses. The virions of archaeal viruses have shapes suggestive of lemons (*Fuselloviridae*), bottles (*Ampullaviridae*), droplets (*Guttaviridae*), or spirals (*Spiraviridae*) (Figure 2; Baquero et al., 2020). Most archaeoviral families encode proteins with no detectable homologs in other viral groups, which hinders their assignment into any of the established realms of the virosphere.

The molecular adaptations of viruses to withstand the extreme extracellular conditions in these habitats remain underexplored, but structural studies have revealed special features in virions isolated from these locations. Some archaeoviruses package their genomes in the A-form of DNA (DiMaio et al., 2015). This DNA conformation is more compact than the more usual B-form, suggesting that this DNA stabilization mechanism is an adaptation to extreme conditions (Wang et al., 2020a). Another archaeovirus, *Aeropyrum coil-shaped virus* (not only the largest known ssDNA virus, but also stable at 95°C), forms a nucleoprotein with its circular ssDNA genome, which coils into an intertwining fiber that compacts adopting the shape of a spiral (Mochizuki et al., 2012). Archaeoviruses can also incorporate ether lipids from their extremophilic host in their virions, which seem to protect against chemical stress (Boyd et al., 2013). Archaeal lipids are also present in the structure of hyperthermophilic spindle-shaped viruses, but its location remains under debate (Quemin et al., 2015; Han and Yuan, 2022).

These structural adaptations to attain virion stability in extreme conditions highlight the ability of viruses to disperse within, but also beyond, the natural habitat of their hosts.

Spindle-shaped archaeoviruses are found in a wide variety of extreme environments around the globe, including deep-sea hydrothermal vents, hypersaline environments, anoxic freshwaters, cold Antarctic lakes, terrestrial hot springs and acidic mines (Wiedenheft et al., 2004; Krupovic et al., 2014), even though their hosts are highly adapted to isolated geographies (Whitaker et al., 2003). However, other studies show that viruses present in hydrothermal vents have restricted host range and are not widely distributed among vent sites (Thomas et al., 2021). All this raises questions about how viruses disperse at a planetary scale (Stedman et al., 2006), and possibly beyond (Griffin, 2013).

## Viral (inter?) planetary distribution

Viruses can disperse passively throughout different geographic scales in their stable virion form. According to the seed-bank hypothesis, viruses can be transported by oceanic or wind currents, and be recruited in ecosystems where they have an opportunity to thrive (Breitbart and Rohwer, 2005). Thus, viruses constitute a global genetic reservoir from which local virus communities reshape within biomes across ecological zones (Brum et al., 2015; Gregory et al., 2019).

Viruses disseminate in the atmosphere associated with dust particles or aerosolized liquid droplets. Therefore, vast amounts of viral particles can be globally spread from events such as desert storms or the release of sea spray (Baylor et al., 1977; Griffin et al., 2001). Some of these virions can be transported past the planetary boundary layer into the troposphere, further favoring long-range viral dispersal (Reche et al., 2018). The deposition of viruses from the troposphere has been observed for air masses coming from both the ocean and the desert, which reflects the magnitude of virus movement across the globe (Reche et al., 2018).

Given the potential of viruses to disperse at a planetary scale, it is likely that some of the viral particles escape the Earth. The same possibility could be applied to extraterrestrial virospheres. In addition to viral adaptations to extreme environments, viruses have been shown –in experimental setups– to survive high doses of UV radiation, ionizing radiation, X rays, high vacuum, desiccation, or microgravity conditions (Koike et al., 1992; Hegedüs et al., 2006; Horneck et al., 2010; Vaishampayan and Grohmann, 2019; Sharma and Curtis, 2022), a fact that would support their possible involvement in the transport of genetic material between planets. Particular consideration deserves the finding that under simulated hot spring conditions, viruses can fossilize upon the aggregation of silica deposits (Laidler and Stedman, 2010; Orange et al., 2011). Since virus silicification renders virions resistant to desiccation, it has been suggested as a potential long-range dispersal mechanism (Laidler et al., 2013). Although the ability of viruses to endure outer space conditions and disperse cosmically remains poorly understood, the field of astrobiology would benefit

from considering viruses as potential carriers of life signatures (Griffin, 2013; Berliner et al., 2018).

## Viruses as biosignatures

The search for extraterrestrial life relies on the study of indicators of extant or past life. Biosignatures are defined as detectable substances, objects or patterns that are likely originated from life processes, and not abiotically (Chan et al., 2019). We could detect extraterrestrial life: (1) remotely, by the observation of astronomical bodies; (2) *in situ*, through space exploration, or (3) in transported samples. A good biosignature should be abundant, unlikely to originate in the absence of life, persistent in time (especially for the detection of past life or in transported samples), and easily detectable (An Astrobiology Strategy for the Search for Life in the Universe, 2019). The expectations of what life looks like in other planets is largely based on our understanding of our own biosphere. However, special attention needs to be put into universal and agnostic biosignatures that could apply to unknown forms of life.

The abundance and importance of viruses in Earth's history and processes suggest that virus-like agents should be abundant members of other biospheres. Would these extraterrestrial molecular parasites form detectable structures like viruses in our planet do?

## Icosahedral capsids as agnostic biosignatures

To increase their stability and protect their genetic material, viruses coat their genomes with a protein capsid encoded in its own genome. However, at least in the life we know, it is impossible to translate a protein large enough to package its own code (i.e., the molecular weight of a codon is  $\sim 1,000$  g/mol and amino acids are 75–203 g/mol, making genetic information bulkier than its product). This size dilemma is overcome in viruses by using capsid proteins with the ability to self-assemble. Thus, a multimeric container can be formed from identical copies of a single gene product. In this case, these protein subunits have the exact chemical and spatial properties and must interact with each other in an equivalent or quasi-equivalent manner, which is achieved by the formation of symmetrical arrangements (Figure 3A; Caspar, 1956; Watson and Crick, 1956). Structural symmetry is a staple of the genetic economy characteristic of viruses, as it allows the creation of containers using the minimal amount of information.

Icosahedral symmetry is a prevalent virus architecture, as the spherical containers it provides are the most efficient in terms of volume-to-surface ratio. Capsids with icosahedral symmetry have emerged multiple times throughout the evolution of the virosphere (Krupovic and Koonin, 2017),

and are present in viruses infecting all domains of life, and packaging any form of nucleic acid (Figure 2). The principle of quasiequivalence (Caspar and Klug, 1962) explains, by a simple rule of triangulation, how any multiple of 60 subunits can assemble into a wide spectrum of structures with icosahedral symmetry. Peptides as short as 24 residues have been shown self-assembly into icosahedral structures (Matsuura et al., 2010), suggesting that these apparently complex architectures could be possible even at early stages of life. Moreover, the icosahedral arrangement of capsid proteins is a direct consequence of a free energy minimization of the chemical interactions between subunits (Zandi et al., 2004), which allows the spontaneous assembly of highly stable structures (Garmann et al., 2019). The geometrical and thermodynamic constraints that pressure the selection of icosahedral symmetries for the formation of molecular containers are probably universal, and as such, in principle should be expected in extraterrestrial virospheres. However, constraints different from those operating in terrestrial life could be relevant in the architecture of the capsids of possible extraterrestrial viruses. When talking about life outside the Earth, it is necessary to take special care that the properties of what we know do not bias our expectations about what we can find. Even in terrestrial viruses, there are other types of capsids apart from the icosahedral ones (Figure 2). An example already cited is the wide variety of morphologies exhibited by archaeal viruses. Another problem that can arise is the detection of structures similar to viral capsids, but whose origin is due to chemical or geological processes. Distinguishing whether an apparently biological structure originates from biotic or abiotic processes is not easy, as has been shown on several occasions, including the controversial detection of microfossils in the ALH84001 meteorite.

Viral structures are good candidates for biosignatures as they are probably abundant components of other biospheres and –given their major role in protecting and dispersing information– can be stable in a wide range of environmental conditions. Despite the small size of viral particles, a plethora of techniques can be used to detect and analyze them. Among them, the most relevant to obtain meaningful structural information are electron microscopy, atomic force microscopy (AFM) and X-ray crystallography. Electron microscopy (cryo-EM in particular) is a powerful method broadly used for structure determination (Jiang and Tang, 2017). However, it requires complex sample manipulation and its throughput is still limited. AFM analyzes the contour of a sample by scanning it mechanically with an ultrafine probe, obtaining virus structure information at high resolution (Figure 3B). AFM technology is advancing at a fast pace, allowing the analysis of several  $\text{mm}^2$  per second at nanometer resolution (Marchesi et al., 2021). AFM devices can be very compact, which is a clear advantage for biosignature detection *in situ*, as they can be –and have been (Smith et al., 2008)– incorporated into spacecrafts. The usage of

X-ray diffraction for the study of virus morphology has seemed to succumb to more convenient techniques such as cryo-EM; however, new advances are revolutionizing the field (Shoemaker and Ando, 2018).

The cutting edge of X-ray diffraction technology is the X-ray free-electron laser (XFEL). The XFEL produces coherent and ultra-bright X-ray femtoseconds-long pulses that can be applied to an aerosolized sample (Chapman et al., 2011). This not only allows for the observation of ultra-fast protein motion at near-atomic resolution (Pandey et al., 2020), it could –most importantly to our matter– scan a sample for molecular structures of interest at a great speed and depth. The XFEL offers a great flexibility in terms of temperature and sample conditions, and it can work under non-equilibrium while penetrating through matter without interfering with external fields (Meents and Wiedorn, 2019). Because icosahedral capsids are highly symmetrical, XFEL can produce clear diffraction patterns from single viral particles with no previous crystallization (Figure 3B; Seibert et al., 2011), which –coupled with machine learning (Suzuki et al., 2020)–, would offer great sensitivity for the detection of icosahedral symmetry in a complex sample. To our knowledge, however, there has not been any astrobiological research efforts in this direction. XFEL technology requires big structures to operate, which would only allow the analysis of samples transported back to Earth; however, technological improvements aimed at reducing the size of XFEL instrumentation are underway (Rosenzweig et al., 2020). Although abiotic molecules with icosahedral symmetry exist –e.g., fullerene and closo-carboranes–, these are much smaller than what is expected from biotic icosahedral structures and should be easily discriminated by the methodologies described above.

## Other viral signatures

Biosignatures based on the detection of chiral excesses, isotopic fractionation or the products of degradation of the biomolecules composing viruses can also be considered as possible proofs of their existence in extraterrestrial biospheres that may even be extinct. Carrying out experiments devoted to determining the chemical and morphological footprint left by viruses immobilized on mineral substrates, before and after being subjected to simulated space conditions in laboratory chambers, can be very useful to know what we can expect to find in another place of space where viruses have been present at some point in its history.

Additionally, there are biosignatures that are not necessarily virus-specific, but that would imply their detection in one way or another. For instance, the usage of chemical polymers to store and transmit information is a likely possibility in other life forms. Given the role of viruses as reservoirs and vehicles of genetic information on Earth and their abundance

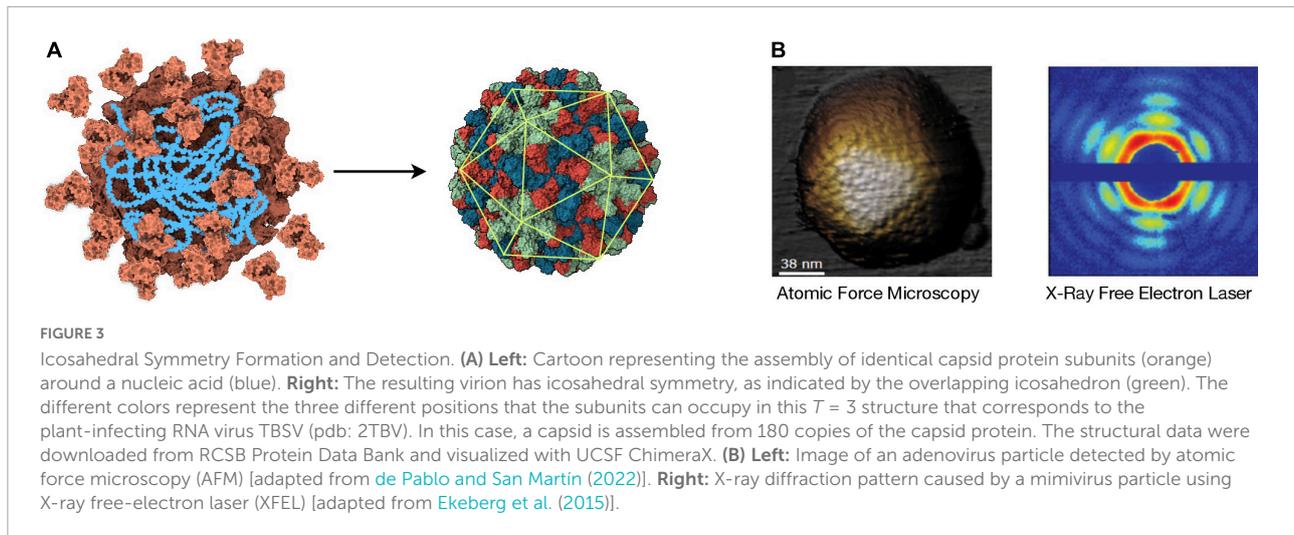
and dispersibility, there are better chances to detect virus-like information packages than more complex life-forms. There are procedures that allow the concentration and purification of viral particles (flocculation with polyethylene glycol and salt, or differential ultracentrifugation) that could be used to increase the probability of life detection in a sample. Sequences technologies based on electrochemistry (e.g., Nanopore) could then be used to detect information patterns from possible biopolymers in a sample, even if they are distinct from DNA or RNA (Rezzonico, 2014). Nanopore-based technology is very compact and, in synergy with other techniques, has successfully been used in samples from a Mars-analog environment, where DNA –including viral sequences– was detected at very low concentrations (Maggiori et al., 2020).

Other life detection methods that are being investigated include immunoassays using antibody microarrays for the recognition of a broad spectrum of common prokaryotic antigens (García-Descalzo et al., 2019). Virus hallmark motifs should be added to this collection of antigens. Antibodies for the detection of the jelly roll fold characteristic of many viral capsids or the ancestral reconstruction of the primitive RRM that antecedes most of viral and cellular replicases (Koonin et al., 2022) would be exceptional candidates for future versions of the LDChip [Life Detector Chip; (García-Descalzo et al., 2019)].

Additional roles of viruses in the detection of life have been discussed (Berliner et al., 2018), including the remote identification of virally provoked phenomena such as the formation of huge calcium carbonate deposits upon lytic termination of algal blooms described in section “Impact of viruses on Earth’s biogeochemical cycles and organismal biodiversity.” Viruses may have also played a role in the transformation of microbial communities, including the generation of microbial mats and stromatolites. If we take into account that the most ancient evidence of life on Earth correspond to stromatolites, the study of whether viruses can manipulate the microbial metabolism to influence carbonate precipitation and other processes related to the lithification of microbial communities can be highly relevant to interpret mineral biosignatures through geological time (White et al., 2021). Research efforts toward the study of viral biomarkers have been scarce, but increasing understanding of the role of viruses in origin and evolution of life will likely put viruses in the spot-light of astrobiological investigations in the near future.

## Viruses as a tool for determining general evolutionary principles

A fundamental topic in Astrobiology is the study of how life has diversified since its origin to give rise to all the forms in which it currently manifests. However, the study of the evolutionary process is not an easy task, because its primary cause –the generation of mutations– occurs by chance and is



subjected to multiple contingencies that condition the action of selective processes. Moreover, its results usually require long periods of time, which makes this process difficult to observe in real time. A valid approach consists of analyzing how biological diversity has changed throughout life history and as a function of the physicochemical parameters of the environment. However, natural environments are complex and are affected by multiple interacting variables whose values cannot be controlled and whose history is frequently unknown. Therefore, relationships between environment, phenotype and genotype are difficult to establish through this strategy, which causes that many basic questions remain unanswered. These include the effect of mutations on fitness, the relative contribution of natural selection and genetic drift, how diverse and reproducible are adaptive pathways, whether error rate is a character subjected to the action of natural selection, or how interactions between mutations and mutants influence adaptation. Experimental evolution arises in this context of uncertainty to provide a framework in which variables can be controlled by the experimenter, allowing a more precise relationship between environmental changes and the response of organisms to them (Bell, 2008; Elena et al., 2008; Kawecki et al., 2012; Elena, 2016; Roux et al., 2016; Geoghegan and Holmes, 2018; Van den Bergh et al., 2018). Populations used in these studies must evolve fast and be easy to handle, two conditions met by many microorganisms, including viruses.

In addition to all exposed above, viruses, and particularly those with an RNA genome, can be a good model for the study of the molecular evolution processes that took place before the emergence of cellular life, when natural selection acted on primitive RNA or RNA-like replicators. RNA viruses, together with viroids and some subgenomic elements, are the only current biological entities that use RNA to store genetic information. This RNA, although

copied by protein enzymes encoded in the viral genome, is actively involved in its own replication, not only by providing binding sites for enzymes and ligands, but also because its copying capacity depends on the three-dimensional structure it adopts according to its primary sequence. Finally, it is widely documented that RNA viruses form population structures in quasispecies, similar to those described theoretically for the ensembles of primitive replicators. All this, together with their lower complexity with respect to cellular systems, means that RNA viruses can be considered a suitable model for the study of evolution in the hypothetical RNA world.

It is impossible to include in this review all the evolutionary questions that viruses have helped to understand. Here we do not intend to make an exhaustive list, and we will simply describe some illustrative examples.

## Effect of mutations on fitness

This question has been the focus of many experimental and theoretical studies (Cuevas et al., 2012; Vale et al., 2012; Acevedo et al., 2014; Bataillon and Bailey, 2014; Minicka et al., 2017). Domingo-Calap and Sanjuán (2011) carried out a comparative study of three phages with a ssRNA genome and three others with a ssDNA genome. Their findings showed that mutations usually have negative effects and confirmed that RNA viruses accumulated mutations at a faster rate than DNA viruses (Sanjuán et al., 2004; Carrasco et al., 2007; Domingo-Calap et al., 2009).

Small genomes, as such of RNA viruses, contain highly compacted information in which mutations are more prone to interact than in larger genomes. This kind of interactions, called epistasis, can make it very difficult to determine the relative contribution of particular mutations to the observed

changes in fitness. Epistatic effects can change depending on the environment, as it was demonstrated in a study carried out with Tobacco Etch Virus that compared the effect of pairs of mutations in different hosts (Cervera et al., 2016). Epistasis may also determine evolutionary trajectories, depending on the first mutations that arise during adaptation (Zhao et al., 2019). A particular case of epistasis are compensatory mutations that are neutral or deleterious mutations that turn beneficial in the context of other deleterious mutations. Their existence is widely documented in drug-resistant mutants where the fitness cost that usually has the resistance mutation is compensated by the acquisition of new mutations (Buckheit, 2004; Boucher et al., 2009).

## Standing genetic diversity vs. diversity generated *de novo*

An evolutionary relevant question is the relative importance for adaptation of the mutations generated *de novo* vs. those pre-existing in populations. Quantitative genetics studies predict that, initially, natural selection acts on the pre-existing genetic variation, while mutations generated *de novo* are more relevant in the long term. The high genetic diversity contained in RNA virus populations provides an excellent system to study this question.

It has been shown that minority genomes, which had selective advantages against previous selective pressures, can speed up adaptation when the population faces a similar condition (Ruíz-Jarabo et al., 2002; Briones and Domingo, 2008). Recent studies carried out with hepatitis C virus showed that the diversification of the mutant spectrum that takes place when it is propagated under constant conditions included the presence of some variants that were resistant to particular antiviral drugs to which the virus had not been previously exposed (Gallego et al., 2020). Other study showed that populations of bacteriophage Q $\beta$  propagated at 37°C contain a considerable fraction of low-frequency mutations that may facilitate adaptation when the virus is exposed to 43°C (Somovilla et al., 2022). The conclusion is that the spread of populations on the space of sequences is unavoidable and play an essential role in adaptation (Koelle et al., 2006; Quakkelaar et al., 2007).

## Influence of error rate on adaptation

High error rates can constitute a great advantage for RNA virus adaptation. However, since most mutations are deleterious, a possible consequence is that small increases in the error rate could lead to the generation of increasingly unfit populations, which could ultimately become extinct. There

are few studies on the effect of environmental conditions on the virus error rate, which would be highly relevant to determine its importance for virus permanence in a particular environment. In contrast, there are many experiments in which the error rate has been artificially modified through the use of mutagens (Domingo et al., 2021). This kind of assays has led to considerable progress in the knowledge of the relationships between error rate, adaptive capacity and the risk of population extinction. Many of the experiments performed in this context were inspired in the concept of error catastrophe of quasispecies theory (Biebricher and Eigen, 2005). It was found that, although in most cases viral populations eventually lost their infectivity when treated with mutagens, no evidence was found of the predicted loss of genetic information. Rather, what seemed to happen was a progressive decrease in fitness due to the increased accumulation of deleterious mutations (Bull et al., 2007).

In some cases, viral populations subjected to mutagenic action were able to select mutants with higher fidelity than the original virus (Pfeiffer and Kirkegaard, 2003; Agudo et al., 2010). Sometimes, these mutations only manifested their effect in the presence of the mutagen (Cabanillas et al., 2014), but in others, the increase in fidelity was maintained regardless of whether the mutagen was present or not (Agudo et al., 2016). These results are highly relevant, since they indicate that the error rate is a character subjected to the action of natural selection, and can be modified. Subsequent experiments carried out with fidelity mutants showed that they had a disadvantage in some infections (Vignuzzi et al., 2006; Coffey et al., 2011), suggesting that the presence of a complex mutant spectrum is necessary for the expression of certain phenotypic traits that require cooperation between different components of the population.

Finally, studies carried out with mutagens have made it possible to identify a new pathway to extinction based on the increase of defective interactions in the mutant spectrum (Grande-Pérez et al., 2005).

## Adaptation to adverse conditions in the extracellular medium

For a virus, as for primitive replicators, the capacity to resist adverse environmental conditions in the inter-replication periods is as important as the capacity to replicate successfully. The ability of viruses to increase their stability in the environment has been explored in several studies (McGee et al., 2014; Costello et al., 2015; Hanson et al., 2015; Lázaro et al., 2018; Whittington and Rokyta, 2019), as well as the possible trade-offs between increased stability and the virus multiplicative capacity (De Paepe and Taddei, 2006; Dessau et al., 2012). A study carried out with vesicular stomatitis virus evolved through a regime that involved an increase in the time that the virus spent out of the host showed the

selection for virus variants with increased extracellular survival and lower fecundity (Elena, 2001; Wasik et al., 2015), which seemed to confirm the trade-off. In contrast to this, other studies carried out with phages – propagated through successive cycles of exposure to adverse extracellular conditions followed by replication at optimal conditions – showed that it was possible to increase the resistance to the harsh extracellular environment without apparent trade-offs on virus replication (McGee et al., 2014; Lázaro et al., 2018).

## Discussion

For the first time in human history, in recent decades we have reached a technological level that allows us to investigate the possible existence of life on other worlds. This is a huge challenge that, if achieved, would help us understand the meaning of life and which of its properties are essential and must be present in any of its manifestations, be they terrestrial or extraterrestrial. The search for life outside of Earth is inevitably biased by the properties of the life we know. Although this fact introduces limitations, it also provides a starting point that can guide us, provided that we are able to stop thinking about the concrete manifestations of life on Earth and start thinking in terms of the processes that life carries out. In this sense, life can be understood as a process that transforms energy using a genetic substrate to store the instructions on how to do it. Viruses do not transform energy by themselves, but they contain in their genome all the necessary instructions to manipulate the metabolism of the cells they infect, so that they manage to multiply and transmit between cells and organisms. Ultimately, the idea of metabolic autonomy is nothing more than an illusion, since, to a greater or lesser degree, most of the living beings on Earth depend on others to exist.

The transition between living and inert matter was probably a continuous process in which it is not clear where to draw the dividing line. In addition to viruses, on Earth there is a set of biological molecules –viroids, transposons, plasmids, prions– that are capable of moving between cells and organisms, producing copies of themselves with a certain independence, and that interact with life in multiple ways, accelerating its evolution and contributing to shaping it. If we want to understand life in its broadest sense, all these entities must be included in the field of study of astrobiology. We do not know if the existence of a biosphere without viruses or similar pathogens is possible, but if this were so, it would almost certainly be less rich and diverse than the one we know. As we have widely described in this review, viruses can help understand the

origin, evolution, distribution, and future of life. Therefore, we encourage the astrobiology community and the funding agencies to give more support to the research on these fascinating entities, as well as to progress in the design of techniques that allow us to include viruses as signatures of the existence of life.

## Author contributions

Both authors designed the structure of this review, wrote the text, carried out the bibliographic search, reviewed the manuscript, and approved the submitted version.

## Funding

This work has been funded by grant PID2020-113284GB-C22, given by the Spanish Ministry of Science and Innovation/State Agency of Research (MCIN/AEI/10.13039/501100011033), and by “ERDF A way of making Europe”. IH was supported by NSF grant MCB-2025305.

## Acknowledgments

We thank our colleagues at the Centro de Astrobiología and the Center for Life in Extreme Environments for the numerous discussions on the topic covered here. We would also like to acknowledge the valuable comments of two reviewers who have undoubtedly contributed to improve the quality of this review.

## Conflict of interest

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Abergel, C., and Claverie, J. M. (2020). Giant viruses. *Curr. Biol.* 30, R1108–R1110. doi: 10.1016/j.cub.2020.08.055
- Abrahão, J., Silva, L., Silva, L. S., Khalil, J. Y. B., Rodrigues, R., Arantes, T., et al. (2018). Tailed giant *Tupanvirus* possesses the most complete translational apparatus of the known virosphere. *Nat. Commun.* 9:749. doi: 10.1038/S41467-018-03168-1
- Abrescia, N. G. A., Bamford, D. H., Grimes, J. M., and Stuart, D. I. (2012). Structure unifies the viral universe. *Annu. Rev. Biochem.* 81, 795–822. doi: 10.1146/annurev-biochem-060910-095130
- Abreu, A., Bourgois, E., Gristwood, A., Troublé, R., Acinas, S. G., Bork, P., et al. (2022). Priorities for ocean microbiome research. *Nat. Microbiol.* 7, 937–947. doi: 10.1038/s41564-022-01145-5
- Acevedo, A., Brodsky, L., and Andino, R. (2014). Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature* 505, 686–690. doi: 10.1038/NATURE12861
- Agudo, R., De la Higuera, I., Arias, A., Grande-Pérez, A., and Domingo, E. (2016). Involvement of a joker mutation in a polymerase-independent lethal mutagenesis escape mechanism. *Virology* 494, 257–266. doi: 10.1016/j.virol.2016.04.023
- Agudo, R., Ferrer-Orta, C., Arias, A., de la Higuera, I., Perales, C., Pérez-Luque, R., et al. (2010). A multi-step process of viral adaptation to a mutagenic nucleoside analogue by modulation of transition types leads to extinction-escape. *PLoS Pathog.* 6:e1001072. doi: 10.1371/journal.ppat.1001072
- Altman, S. (1989). Ribonuclease P: An enzyme with a catalytic RNA subunit. *Adv. Enzymol. Relat. Areas Mol. Biol.* 62, 1–36. doi: 10.1002/9780470123089.CH1
- An Astrobiology Strategy for the Search for Life in the Universe. (2019). *An astrobiology strategy for the search for life in the universe*. Washington, DC: National Academies Press. doi: 10.17226/25252
- Anantharaman, K., Duhaime, M. B., Breier, J. A., Wendt, K. A., Toner, B. M., and Dick, G. J. (2014). Sulfur oxidation genes in diverse deep-sea viruses. *Science* 344, 757–760. doi: 10.1126/science.1252229
- Ankrah, N. Y. D., May, A. L., Middleton, J. L., Jones, D. R., Hadden, M. K., Gooding, J. R., et al. (2014). Phage infection of an environmentally relevant marine bacterium alters host metabolism and lysate composition. *ISME J.* 8, 1089–1100. doi: 10.1038/ismej.2013.216
- Arnaud, F., Varela, M., Spencer, T. E., and Palmerini, M. (2008). Coevolution of endogenous betaretroviruses of sheep and their host. *Cell. Mol. Life Sci.* 65, 3422–3432. doi: 10.1007/S00018-008-8500-9
- Baaske, P., Weinert, F. M., Dühr, S., Lemke, K. H., Russell, M. J., and Braun, D. (2007). Extreme accumulation of nucleotides in simulated hydrothermal pore systems. *Proc. Natl. Acad. Sci. U.S.A.* 104, 9346–9351. doi: 10.1073/pnas.0609521104
- Bai, J., Yang, Z., Li, H., Hong, Y., Fan, D., Lin, A., et al. (2021). Genome-wide characterization of zebrafish endogenous retroviruses reveals unexpected diversity in genetic organizations and functional potentials. *Microbiol. Spectr.* 9:e0225421. doi: 10.1128/spectrum.02254-21
- Baltimore, D. (1971). Expression of animal virus genomes. *Bacteriol. Rev.* 35, 235–241. doi: 10.1128/MMBR.35.3.235-241.1971
- Bamford, D. H. (2003). Do viruses form lineages across different domains of life? *Res. Microbiol.* 154, 231–236. doi: 10.1016/S0923-2508(03)00065-2
- Baquero, D. P., Liu, Y., Wang, F., Egelman, E. H., Prangishvili, D., and Krupovic, M. (2020). *Structure and assembly of archaeal viruses*, 1st Edn. Amsterdam: Elsevier Inc. doi: 10.1016/bs.aivir.2020.09.004
- Barnard, C. J. (1984). Stasis: A coevolutionary model. *J. Theor. Biol.* 110, 27–34. doi: 10.1016/S0022-5193(84)80013-2
- Bass, B. L., and Cech, T. R. (1984). Specific interaction between the self-splicing RNA of *Tetrahymena* and its guanosine substrate: Implications for biological catalysis by RNA. *Nature* 308, 820–826. doi: 10.1038/308820a0
- Bataillon, T., and Bailey, S. F. (2014). Effects of new mutations on fitness: Insights from models and data. *Ann. N.Y. Acad. Sci.* 1320, 76–92. doi: 10.1111/nyas.12460
- Batschelet, E., Domingo, E., and Weissmann, C. (1976). The proportion of revertant and mutant phage in a growing population, as a function of mutation and growth rate. *Gene* 1, 27–32. doi: 10.1016/0378-1119(76)90004-4
- Baylor, E. R., Baylor, M. B., Blanchard, D. C., Syzdek, L. D., and Appel, C. (1977). Virus transfer from surf to wind. *Science* 198, 575–580. doi: 10.1126/science.918656
- Bell, G. (2008). Experimental evolution. *Heredity (Edinb.)* 100, 441–442. doi: 10.1038/hdy.2008.19
- Bell, P. J. L. (2020). Evidence supporting a viral origin of the eukaryotic nucleus. *Virus Res.* 289:198168. doi: 10.1016/J.VIRUSRES.2020.198168
- Benner, S. A. (2010). Defining life. *Astrobiology* 10, 1021–1030. doi: 10.1089/AST.2010.0524
- Bennett, J. O., Shostak, G. S., Schneider, N., and MacGregor, M. (2022). *Life in the universe*, 5th Edn. Princeton, NJ: Princeton University Press.
- Benson, S. D., Bamford, J. K. H., Bamford, D. H., and Burnett, R. M. (2004). Does common architecture reveal a viral lineage spanning all three domains of life? *Mol. Cell* 16, 673–685. doi: 10.1016/J.MOLCEL.2004.11.016
- Bergh, O., Børsheim, K. Y., Bratbak, G., and Heldal, M. (1989). High abundance of viruses found in aquatic environments. *Nature* 340, 467–468. doi: 10.1038/340467a0
- Berliner, A. J., Mochizuki, T., and Steadman, K. M. (2018). Astrovirology: Viruses at large in the universe. *Astrobiology* 18, 207–223. doi: 10.1089/ast.2017.1649
- Biebricher, C. K., and Eigen, M. (2005). The error threshold. *Virus Res.* 107, 117–127. doi: 10.1016/J.VIRUSRES.2004.11.002
- Biebricher, C. K., and Eigen, M. (2006). “What is a quasispecies?” in *Quasispecies: Concept and implications for virology. Current topics in microbiology and immunology*, ed. E. Domingo (Berlin: Springer-Verlag).
- Biebricher, C. K., and Gardiner, W. C. (1997). Molecular evolution of RNA in vitro. *Biophys. Chem.* 66, 179–192. doi: 10.1016/S0301-4622(97)00059-8
- Boucher, C. A. B., Nijhuis, M., and Van Maarseveen, N. M. (2009). Antiviral resistance and impact on viral replication capacity: Evolution of viruses under antiviral pressure occurs in three phases. *Handb. Exp. Pharmacol.* 189, 299–320. doi: 10.1007/978-3-540-79086-0\_11
- Boyd, E. S., Hamilton, T. L., Wang, J., He, L., and Zhang, C. L. (2013). The role of tetraether lipid composition in the adaptation of thermophilic archaea to acidity. *Front. Microbiol.* 4:62. doi: 10.3389/fmicb.2013.00062
- Brahim Belhaouari, D., Pires De Souza, G. A., Lamb, D. C., Kelly, S. L., Goldstone, J. V., Stegeman, J. J., et al. (2022). Metabolic arsenal of giant viruses: Host hijack or self-use? *Elife* 11:e78674. doi: 10.7554/eLife.78674
- Breitbart, M., and Rohwer, F. (2005). Here a virus, there a virus, everywhere the same virus? *Trends Microbiol.* 13, 278–284. doi: 10.1016/j.tim.2005.04.003
- Briones, C., and Domingo, E. (2008). Minority report: Hidden memory genomes in HIV-1 quasispecies and possible clinical implications. *AIDS Rev.* 10, 93–109.
- Brum, J. R., Ignacio-Espinoza, J. C., Roux, S., Doulcier, G., Acinas, S. G., Alberti, A., et al. (2015). Patterns and ecological drivers of ocean viral communities. *Science* 348:1261498. doi: 10.1126/science.1261498
- Brum, J. R., Schenck, R. O., and Sullivan, M. B. (2013). Global morphological analysis of marine viruses shows minimal regional variation and dominance of non-tailed viruses. *ISME J.* 7, 1738–1751. doi: 10.1038/ISMEJ.2013.67
- Brussaard, C. P. D. (2004). Viral control of phytoplankton populations—a review. *J. Eukaryot. Microbiol.* 51, 125–138. doi: 10.1111/j.1550-7408.2004.tb00537.x
- Buckheit, R. W. (2004). Understanding HIV resistance, fitness, replication capacity and compensation: Targeting viral fitness as a therapeutic strategy. *Expert Opin. Investig. Drugs* 13, 933–958. doi: 10.1517/13543784.13.8.933
- Buckling, A., and Brockhurst, M. (2012). Bacteria-virus coevolution. *Adv. Exp. Med. Biol.* 751, 347–370. doi: 10.1007/978-1-4614-3567-9\_16
- Bull, J. J., Sanjuán, R., and Wilke, C. O. (2007). Theory of lethal mutagenesis for viruses. *J. Virol.* 81, 2930–2939. doi: 10.1128/JVI.01624-06
- Cabanillas, L., Sanjuán, R., and Lázaro, E. (2014). Changes in protein domains outside the catalytic site of the bacteriophage Q $\beta$  replicase reduce the mutagenic effect of 5-azacytidine. *J. Virol.* 88, 10480–10487. doi: 10.1128/JVI.00979-14
- Carrasco, P., de la Iglesia, F., and Elena, S. F. (2007). Distribution of fitness and virulence effects caused by single-nucleotide substitutions in Tobacco Etch virus. *J. Virol.* 81, 12979–12984. doi: 10.1128/JVI.00524-07
- Caspar, D. L. D. (1956). Structure of bushy stunt virus. *Nature* 177, 475–476. doi: 10.1038/177475a0
- Caspar, D. L. D., and Klug, A. (1962). Physical principles in the construction of regular viruses. *Cold Spring Harb. Symp. Quant. Biol.* 27, 1–24. doi: 10.1101/SQB.1962.027.001.005
- Castelán-Sánchez, H. G., López-Rosas, I., García-Suastegui, W. A., Peralta, R., Dobson, A. D. W., Batista-García, R. A., et al. (2019). Extremophile deep-sea viral

- communities from hydrothermal vents: Structural and functional analysis. *Mar. Genomics* 46, 16–28. doi: 10.1016/j.margen.2019.03.001
- Cervera, H., Lalić, J., and Elena, S. F. (2016). Effect of host species on topography of the fitness landscape for a plant RNA virus. *J. Virol.* 90, 10160–10169. doi: 10.1128/JVI.01243-16
- Chaikerasitak, V., Nguyen, K., Khanna, K., Brilot, A. F., Erb, M. L., Coker, J. K. C., et al. (2017). Assembly of a nucleus-like structure during viral replication in bacteria. *Science* 355, 194–197. doi: 10.1126/SCIENCE.AAL2130
- Chan, M. A., Hinman, N. W., Potter-McIntyre, S. L., Schubert, K. E., Gillams, R. J., Awramik, S. M., et al. (2019). Deciphering biosignatures in planetary contexts. *Astrobiology* 19, 1075–1102. doi: 10.1089/ast.2018.1903
- Chapman, H. N., Fromme, P., Barty, A., White, T. A., Kirian, R. A., Aquila, A., et al. (2011). Femtosecond X-ray protein nanocrystallography. *Nature* 470, 73–78. doi: 10.1038/nature09750
- Chapman, K. B., and Szostak, J. W. (1994). In vitro selection of catalytic RNAs. *Curr. Opin. Struct. Biol.* 4, 618–622. doi: 10.1016/S0959-440X(94)90227-5
- Cleland, C. E. (2011). Life without definitions. *Synthese* 185(1), 125–144. doi: 10.1007/S11229-011-9879-7
- Cleland, C. E. (2019). Moving beyond definitions in the search for extraterrestrial life. *Astrobiology* 19, 722–729. doi: 10.1089/ast.2018.1980
- Cléry, A., Blatter, M., and Allain, F. H. T. (2008). RNA recognition motifs: Boring? Not quite. *Curr. Opin. Struct. Biol.* 18, 290–298. doi: 10.1016/J.SBI.2008.04.002
- Cockell, C. (2020). *Astrobiology?: Understanding life in the universe*, 2nd Edn. New York, NY: Wiley-Blackwell.
- Coffey, L. L., Beechey, Y., Borderia, A. V., Blanc, H., and Vignuzzi, M. (2011). Arbovirus high fidelity variant loses fitness in mosquitoes and mice. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16038–16043. doi: 10.1073/PNAS.1111650108
- Cole, C. N. (1975). Defective interfering (DI) particles of poliovirus. *Prog. Med. Virol.* Vol. 20, 180–207.
- Costello, D. A., Whittaker, G. R., and Daniel, S. (2015). Variations in pH sensitivity, acid stability, and fusogenicity of three influenza virus H3 subtypes. *J. Virol.* 89, 350–360. doi: 10.1128/JVI.01927-14
- Cuevas, J. M., Domingo-Calap, P., and Sanjuán, R. (2012). The fitness effects of synonymous mutations in DNA and RNA viruses. *Mol. Biol. Evol.* 29, 17–20. doi: 10.1093/molbev/msr179
- Daly, R. A., Roux, S., Borton, M. A., Morgan, D. M., Johnston, M. D., Booker, A. E., et al. (2019). Viruses control dominant bacteria colonizing the terrestrial deep biosphere after hydraulic fracturing. *Nat. Microbiol.* 4, 352–361. doi: 10.1038/S41564-018-0312-6
- de la Higuera, I., Kasun, G. W., Torrance, E. L., Pratt, A. A., Maluenda, A., Colombet, J., et al. (2020). Unveiling cruciviral diversity by mining metagenomic data. *mBio* 11, e1410–e1420. doi: 10.1128/mBio.01410-20
- de Pablo, P. J., and San Martín, C. (2022). Seeing and touching adenovirus: Complementary approaches for understanding assembly and disassembly of a complex virion. *Curr. Opin. Virol.* 52, 112–122. doi: 10.1016/j.coviro.2021.11.006
- De Paepe, M., and Taddei, F. (2006). Viruses' life history: Towards a mechanistic basis of a trade-off between survival and reproduction among phages. *PLoS Biol.* 4:e193. doi: 10.1371/JOURNAL.PBIO.0040193
- Des Marais, D. J., Nuth, J. A., Allamandola, L. J., Boss, A. P., Farmer, J. D., Hoehler, T. M., et al. (2008). The NASA astrobiology roadmap. *Astrobiology* 8, 715–730. doi: 10.1089/ast.2008.0819
- Dessau, M., Goldhill, D., McBride, R. L., Turner, P. E., and Modis, Y. (2012). Selective pressure causes an RNA virus to trade reproductive fitness for increased structural and thermal stability of a viral enzyme. *PLoS Genet.* 8:e1003102. doi: 10.1371/JOURNAL.PGEN.1003102
- DiMaio, F., Yu, X., Rensen, E., Krupovic, M., Prangishvili, D., and Egelman, E. H. (2015). A virus that infects a hyperthermophile encapsidates A-form DNA. *Science* 348, 914–917. doi: 10.1126/science.aaa4181
- Dix, D. E. (2003). What is life? Prerequisites for a definition. *Yale J. Biol. Med.* 75, 313–321.
- Domingo, E., García-Crespo, C., Lobo-Vega, R., and Perales, C. (2021). Mutation rates, mutation frequencies, and proofreading-repair activities in rna virus genetics. *Viruses* 13:1882. doi: 10.3390/v13091882
- Domingo, E., Sabo, D., Taniguchi, T., and Weissmann, C. (1978). Nucleotide sequence heterogeneity of an RNA phage population. *Cell* 13, 735–744. doi: 10.1016/0092-8674(78)90223-4
- Domingo-Calap, P., and Sanjuán, R. (2011). Experimental evolution of rna versus dna viruses. *Evolution* 65, 2987–2994. doi: 10.1111/j.1558-5646.2011.01339.x
- Domingo-Calap, P., Cuevas, J. M., and Sanjuán, R. (2009). The fitness effects of random mutations in single-stranded DNA and RNA bacteriophages. *PLoS Genet.* 5:e1000742. doi: 10.1371/journal.pgen.1000742
- Dominguez-Huerta, G., Zayed, A. A., Wainaina, J. M., Guo, J., Tian, F., Pratama, A. A., et al. (2022). Diversity and ecological footprint of global ocean RNA viruses. *Science* 376, 1202–1208. doi: 10.1126/science.abn6358
- Duffy, S., Shackelton, L. A., and Holmes, E. C. (2008). Rates of evolutionary change in viruses: Patterns and determinants. *Nat. Rev. Genet.* 9, 267–276. doi: 10.1038/nrg2323
- Dupressoir, A., Vernochet, C., Harper, F., Guégan, J., Dessen, P., Pierron, G., et al. (2011). A pair of co-opted retroviral envelope syncytin genes is required for formation of the two-layered murine placental syncytiotrophoblast. *Proc. Natl. Acad. Sci. U.S.A.* 108, E1164–E1173. doi: 10.1073/PNAS.1112304108
- Durzyńska, J., and Goździcka-Józefiak, A. (2015). Viruses and cells intertwined since the dawn of evolution. *Virol. J.* 12:169. doi: 10.1186/s12985-015-0400-7
- Edgar, R. C., Taylor, J., Lin, V., Altman, T., Barbera, P., Meleshko, D., et al. (2022). Petabase-scale sequence alignment catalyses viral discovery. *Nature* 602, 142–147. doi: 10.1038/s41586-021-04332-2
- Eigen, M. (1971). Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 58, 465–523.
- Eigen, M., and Schuster, P. (1979). *The hypercycle a principle of natural self-organization*. Berlin: Springer.
- Ekeberg, T., Svenda, M., Abergel, C., Maia, F. R. N. C., Seltzer, V., Claverie, J. M., et al. (2015). Three-dimensional reconstruction of the giant mimivirus particle with an X-ray free-electron laser. *Phys. Rev. Lett.* 114:098102. doi: 10.1103/PHYSREVLETT.114.098102
- Elena, S. F. (2001). Evolutionary history conditions the timing of transmission in vesicular stomatitis virus. *Infect. Genet. Evol.* 1, 151–159. doi: 10.1016/S1567-1348(01)00022-3
- Elena, S. F. (2016). Evolutionary transitions during RNA virus experimental evolution. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 371:20150441. doi: 10.1098/RSTB.2015.0441
- Elena, S. F., Agudelo-Romero, P., Carrasco, P., Codoñer, F. M., Martín, S., Torres-Barceló, C., et al. (2008). Experimental evolution of plant RNA viruses. *Heredity (Edinb)* 100, 478–483. doi: 10.1038/sj.hdy.6801088
- Emerson, J. B., Roux, S., Brum, J. R., Bolduc, B., Woodcroft, B. J., Jang, H. B., et al. (2018). Host-linked soil viral ecology along a permafrost thaw gradient. *Nat. Microbiol.* 3, 870–880. doi: 10.1038/s41564-018-0190-y
- Feiner, R., Argov, T., Rabinovich, L., Sigal, N., Borovok, I., and Herskovits, A. A. (2015). A new perspective on lysogeny: Prophages as active regulatory switches of bacteria. *Nat. Rev. Microbiol.* 13(10), 641–650. doi: 10.1038/nrmicro3527
- Feschotte, C. (2008). Transposable elements and the evolution of regulatory networks. *Nat. Rev. Genet.* 9, 397–405. doi: 10.1038/NRG2337
- Feschotte, C., and Gilbert, C. (2012). Endogenous viruses: Insights into viral evolution and impact on host biology. *Nat. Rev. Genet.* 13, 283–296. doi: 10.1038/NRG3199
- Filée, J., and Forterre, P. (2005). Viral proteins functioning in organelles: A cryptic origin? *Trends Microbiol.* 13, 510–513. doi: 10.1016/J.TIM.2005.08.012
- Filée, J., Forterre, P., and Laurent, J. (2003). The role played by viruses in the evolution of their hosts: A view based on informational protein phylogenies. *Res. Microbiol.* 154, 237–243. doi: 10.1016/S0923-2508(03)00066-4
- Fischer, M. G. (2011). Sputnik and mavirus: More than just satellite viruses. *Nat. Rev. Microbiol.* 10:88. doi: 10.1038/NRMICRO2676-C1
- Flombaum, P., Gallegos, J. L., Gordillo, R. A., Rincón, J., Zabala, L. L., Jiao, N., et al. (2013). Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 9824–9829. doi: 10.1073/pnas.1307701110
- Forterre, P. (2006). The origin of viruses and their possible roles in major evolutionary transitions. *Virus Res.* 117, 5–16. doi: 10.1016/J.VIRUSRES.2006.01.010
- Forterre, P. (2011). Manipulation of cellular syntheses and the nature of viruses: The virocell concept. *Comptes Rendus Chim.* 14, 392–399. doi: 10.1016/J.CRCI.2010.06.007
- Forterre, P. (2016). To be or not to be alive: How recent discoveries challenge the traditional definitions of viruses and life. *Stud. Hist. Philos. Biol. Biomed. Sci.* 59, 100–108. doi: 10.1016/j.shpsc.2016.02.013
- Forterre, P., and Prangishvili, D. (2009). The origin of viruses. *Res. Microbiol.* 160, 466–472. doi: 10.1016/J.RESMIC.2009.07.008

- French, R. K., and Holmes, E. C. (2020). An ecosystems perspective on virus evolution and emergence. *Trends Microbiol.* 28, 165–175. doi: 10.1016/J.TIM.2019.10.010
- Fridman, S., Flores-Urbe, J., Larom, S., Alalouf, O., Liran, O., Yacoby, I., et al. (2017). A myovirus encoding both photosystem I and II proteins enhances cyclic electron flow in infected *Prochlorococcus* cells. *Nat. Microbiol.* 2, 1350–1357. doi: 10.1038/s41564-017-0002-9
- Fridmann-Sirkis, Y., Milrot, E., Mutsafi, Y., Ben-Dor, S., Levin, Y., Savidor, A., et al. (2016). Efficiency in complexity: Composition and dynamic nature of mimivirus replication factories. *J. Virol.* 90, 10039–10047. doi: 10.1128/JVI.01319-16
- Gabashvili, E., Kobakhidze, S., Chkhikvishvili, T., Tabatadze, L., Tsiklauri, R., Dadiani, K., et al. (2022). Bacteriophage-mediated risk pathways underlying the emergence of antimicrobial resistance via intragenetic and intergeneric recombination of antibiotic efflux genes across natural populations of human pathogenic bacteria. *Microb. Ecol.* 84, 213–226. doi: 10.1007/s00248-021-01846-0
- Gabashvili, E., Osepashvili, M., Koulouris, S., Ujmajuridze, L., Tskhishvili, Z., and Kotetishvili, M. (2020). Phage transduction is involved in the intergeneric spread of antibiotic resistance-associated bla CTX-M, mel, and tetM loci in natural populations of some human and animal bacterial pathogens. *Curr. Microbiol.* 77, 185–193. doi: 10.1007/s00284-019-01817-2
- Gago, S., Elena, S. F., Flores, R., and Sanjuán, R. (2009). Extremely high mutation rate of a hammerhead viroid. *Science* 323:1308. doi: 10.1126/SCIENCE.1169202
- Gallego, I., Soria, M. E., García-Crespo, C., Chen, Q., Martínez-Barragán, P., Khalifaoui, S., et al. (2020). Broad and dynamic diversification of infectious hepatitis C virus in a cell culture environment. *J. Virol.* 94, e01856–19. doi: 10.1128/JVI.01856-19
- Gao, L., Altae-Tran, H., Böhning, F., Makarova, K. S., Segel, M., Schmid-Burgk, J. L., et al. (2020). Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science* 369, 1077–1084. doi: 10.1126/SCIENCE.ABA0372
- García-Descalzo, L., Parro, V., García-Villadangos, M., Cockell, C. S., Moissl-Eichinger, C., Perras, A., et al. (2019). Microbial markers profile in anaerobic mars analogue environments using the lDchip (Life detector chip) antibody microarray core of the solid (signs of life detector) platform. *Microorganisms* 7:365. doi: 10.3390/microorganisms7090365
- Garman, R. F., Goldfain, A. M., and Manoharan, V. N. (2019). Measurements of the self-assembly kinetics of individual viral capsids around their RNA genome. *Proc. Natl. Acad. Sci.* 116, 22485–22490. doi: 10.1073/pnas.190922.116
- Geoghegan, J. L., and Holmes, E. C. (2018). Evolutionary virology at 40. *Genetics* 210, 1151–1162. doi: 10.1534/GENETICS.118.301556
- Gil, J. F., Mesa, V., Estrada-Ortiz, N., Lopez-Obando, M., Gómez, A., and Plácido, J. (2021). Viruses in extreme environments, current overview, and biotechnological potential. *Viruses* 13, 1–23. doi: 10.3390/v13010081
- Gilbert, C., and Cordaux, R. (2017). Viruses as vectors of horizontal transfer of genetic material in eukaryotes. *Curr. Opin. Virol.* 25, 16–22. doi: 10.1016/J.COVIRO.2017.06.005
- Gilbert, W. (1986). Origin of life: The RNA world. *Nature* 319:618. doi: 10.1038/319618a0
- Giovannoni, S., Temperton, B., and Zhao, Y. (2013). Giovannoni et al. reply to “SAR11 viruses and defensive host strains.”. *Nature* 499, E4–E5. doi: 10.1038/nature12388
- Gobler, C. J., Hutchins, D. A., Fisher, N. S., Cosper, E. M., and Sañudo-Wilhelmy, S. A. (1997). Release and bioavailability of C, N, P, Se, and Fe following viral lysis of a marine chrysophyte. *Limnol. Oceanogr.* 42, 1492–1504. doi: 10.4319/lo.1997.42.7.1492
- Grande-Pérez, A., Lázaro, E., Lowenstein, P., Domingo, E., and Manrubia, S. C. (2005). Suppression of viral infectivity through lethal defection. *Proc. Natl. Acad. Sci. U.S.A.* 102:4448. doi: 10.1073/PNAS.0408871102
- Greenwood, A. D., Ishida, Y., O'Brien, S. P., Roca, A. L., and Eiden, M. V. (2018). Transmission, Evolution, and endogenization: Lessons learned from recent retroviral invasions. *Microbiol. Mol. Biol. Rev.* 82, e00044–17. doi: 10.1128/mmb.00044-17
- Gregory, A. C., Zayed, A. A., Conceição-Neto, N., Temperton, B., Bolduc, B., Alberti, A., et al. (2019). Marine DNA viral macro- and microdiversity from pole to pole. *Cell* 177, 1109.e–1123.e. doi: 10.1016/j.cell.2019.03.040
- Griffin, D. W. (2013). The quest for extraterrestrial life: What about the viruses? *Astrobiology* 13, 774–783. doi: 10.1089/AST.2012.0959
- Griffin, D. W., Garrison, V. H., Herman, J. R., and Shinn, E. A. (2001). African desert dust in the Caribbean atmosphere: Microbiology and public health. *Aerobiologia (Bologna)* 17, 203–213. doi: 10.1023/A:1011868218901
- Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., et al. (2016). Europe PMC funders group plankton networks driving carbon export in the oligotrophic ocean. *Nature* 532, 465–470. doi: 10.1038/nature16942
- Hager, A. J., Pollard, J. D., and Szostak, J. W. (1996). Ribozymes: Aiming at RNA replication and protein synthesis. *Chem. Biol.* 3, 717–725. doi: 10.1016/S1074-5521(96)90246-X
- Hampton, H. G., Watson, B. N. J., and Fineran, P. C. (2020). The arms race between bacteria and their phage foes. *Nature* 577, 327–336. doi: 10.1038/S41586-019-1894-8
- Han, Z., and Yuan, W. (2022). Structural insights into a spindle-shaped archaeal virus with a sevenfold symmetrical tail. *Proc. Natl. Acad. Sci. U.S.A.* 119:e2119439119. doi: 10.1073/pnas.2119439119
- Hanson, A., Imai, M., Hatta, M., McBride, R., Imai, H., Taft, A., et al. (2015). Identification of stabilizing mutations in an H5 hemagglutinin influenza virus protein. *J. Virol.* 90, 2981–2992. doi: 10.1128/JVI.02790-15
- Harris, H. M. B., and Hill, C. (2020). A place for viruses on the tree of life. *Front. Microbiol.* 11:604048. doi: 10.3389/FMICB.2020.604048
- Hays, L., Achenbach, L., Bailey, J., Barnes, R., Baross, J., Bertka, C., et al. (2015). *NASA astrobiology strategy 2015*. Washington, DC: National Aeronautics and Space Administration (NASA).
- Hayward, A., Cornwallis, C. K., and Jern, P. (2015). Pan-vertebrate comparative genomics unmasks retrovirus macroevolution. *Proc. Natl. Acad. Sci. U.S.A.* 112, 464–469. doi: 10.1073/PNAS.1414980112
- Hegde, N. R., Maddur, M. S., Kaveri, S. V., and Bayry, J. (2009). Reasons to include viruses in the tree of life. *Nat. Rev. Microbiol.* 7:615. doi: 10.1038/NRMICRO2108-C1
- Hegedüs, M., Kovács, G., Módos, K., Rontó, G., Lammer, H., Panitz, C., et al. (2006). Exposure of phage T7 to simulated space environment: The effect of vacuum and UV-C radiation. *J. Photochem. Photobiol. B.* 82, 94–104. doi: 10.1016/J.JPHOTOBIOL.2005.09.002
- Hendrix, R. W., Smith, M. C. M., Burns, R. N., Ford, M. E., and Hatfull, G. F. (1999). Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. *Proc. Natl. Acad. Sci.* 96, 2192–2197. doi: 10.1073/pnas.96.5.2192
- Higgs, P. G., and Lehman, N. (2015). The RNA world: Molecular cooperation at the origins of life. *Nat. Rev. Genet.* 16, 7–17. doi: 10.1038/NRG3841
- Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., and VandePol, S. (1982). Rapid evolution of RNA genomes. *Science* 215, 1577–1585. doi: 10.1126/SCIENCE.7041255
- Holmes, E. C. (2004). Adaptation and immunity. *PLoS Biol.* 2:E307. doi: 10.1371/JOURNAL.PBIO.0020307
- Holmes, E. C. (2011b). What does virus evolution tell us about virus origins? *J. Virol.* 85, 5247–5251. doi: 10.1128/JVI.02203-10
- Holmes, E. C. (2011a). The evolution of endogenous viral elements. *Cell Host Microbe* 10, 368–377. doi: 10.1016/J.CHOM.2011.09.002
- Holmes, E. C., and Duchêne, S. (2019). Evolutionary stasis of viruses? *Nat. Rev. Microbiol.* 17:329. doi: 10.1038/S41579-019-0168-7
- Horneck, G., Klaus, D. M., and Mancinelli, R. L. (2010). Space microbiology. *Microbiol. Mol. Biol. Rev.* 74, 121–156. doi: 10.1128/MMBR.00016-09
- Howard-Varona, C., Lindback, M. M., Bastien, G. E., Solonenko, N., Zayed, A. A., Jang, H. B., et al. (2020). Phage-specific metabolic reprogramming of virocells. *ISME J.* 14, 881–895. doi: 10.1038/s41396-019-0580-z
- Huang, A. S. (1973). Defective interfering viruses. *Annu. Rev. Microbiol.* 27, 101–117. doi: 10.1146/ANNUREV.MI.27.100173.000533
- Hughes, J. F., and Coffin, J. M. (2001). Evidence for genomic rearrangements mediated by human endogenous retroviruses during primate evolution. *Nat. Genet.* 29, 487–489. doi: 10.1038/NG775
- Hurwitz, B. L., Hallam, S. J., and Sullivan, M. B. (2013). Metabolic reprogramming by viruses in the sunlit and dark ocean. *Genome Biol.* 14:R123. doi: 10.1186/gb-2013-14-11-r123
- Iranzo, J., Puigbo, P., Lobkovsky, A. E., Wolf, Y. I., and Koonin, E. V. (2016). Inevitability of genetic parasites. *Genome Biol. Evol.* 8, 2856–2869. doi: 10.1093/gbe/evw193
- Irwin, N. A. T., Pittis, A. A., Richards, T. A., and Keeling, P. J. (2022). Systematic evaluation of horizontal gene transfer between eukaryotes and viruses. *Nat. Microbiol.* 7, 327–336. doi: 10.1038/s41564-021-01026-3

- Iyer, L. M., Koonin, E. V., Leipe, D. D., and Aravind, L. (2005). Origin and evolution of the archaeo-eukaryotic primase superfamily and related palm-domain proteins: Structural insights and new members. *Nucleic Acids Res.* 33, 3875–3896. doi: 10.1093/NAR/GKI702
- Janzen, E., Janzen, E., Blanco, C., Peng, H., Kenchel, J., Kenchel, J., et al. (2020). Promiscuous ribozymes and their proposed role in prebiotic evolution. *Chem. Rev.* 120, 4879–4897. doi: 10.1021/acs.chemrev.9b00620
- Jiang, W., and Tang, L. (2017). Atomic Cryo-EM structures of viruses. *Curr. Opin. Struct. Biol.* 46:122. doi: 10.1016/j.SBI.2017.07.002
- Joyce, G. F., and Szostak, J. W. (2018). Protocells and RNA Self-Replication. *Cold Spring Harb. Perspect. Biol.* 10:a034801. doi: 10.1101/CSHPERSPECT.A034801
- Kacian, D. L., Mills, D. R., Kramer, F. R., and Spiegelman, S. (1972). A replicating RNA molecule suitable for a detailed analysis of extracellular evolution and replication. *Proc. Natl. Acad. Sci. U.S.A.* 69, 3038–3042. doi: 10.1073/pnas.69.10.3038
- Kaján, G. L., Doszpoly, A., Tarján, Z. L., Vidovszky, M. Z., and Papp, T. (2020). Virus-host coevolution with a focus on animal and human DNA viruses. *J. Mol. Evol.* 88, 41–56. doi: 10.1007/S00239-019-09913-4
- Kaneko, H., Blanc-Mathieu, R., Endo, H., Chaffron, S., Delmont, T. O., Gaia, M., et al. (2021). Eukaryotic virus composition can predict the efficiency of carbon export in the global ocean. *iScience* 24:102002. doi: 10.1016/j.isci.2020.102002
- Kawecki, T. J., Lenski, R. E., Ebert, D., Hollis, B., Olivieri, I., and Whitlock, M. C. (2012). Experimental evolution. *Trends Ecol. Evol.* 27, 547–560. doi: 10.1016/j.TREE.2012.06.001
- Kazlauskas, D., Krupovic, M., and Venclovas, C. (2016). The logic of DNA replication in double-stranded DNA viruses: Insights from global analysis of viral genomes. *Nucleic Acids Res.* 44, 4551–4564. doi: 10.1093/NAR/GKW322
- Kazlauskas, D., Varsani, A., Koonin, E. V., and Krupovic, M. (2019). Multiple origins of prokaryotic and eukaryotic single-stranded DNA viruses from bacterial and archaeal plasmids. *Nat. Commun.* 10, 1–12. doi: 10.1038/s41467-019-11433-0
- Koelle, K., Cobey, S., Grenfell, B., and Pascual, M. (2006). Epochal evolution shapes the phylodynamics of inter-pandemic influenza A (H3N2) in humans. *Science* 314, 1898–1903. doi: 10.1126/SCIENCE.1132745
- Koike, J., Oshima, T., Koike, K. A., Taguchi, H., Tanaka, R., Nishimura, K., et al. (1992). Survival rates of some terrestrial microorganisms under simulated space conditions. *Adv. Sp. Res.* 12, 271–274. doi: 10.1016/0273-1177(92)90182-W
- Koonin, E. V. (2014). The origins of cellular life. *Antonie Van Leeuwenhoek* 106, 27–41. doi: 10.1007/S10482-014-0169-5
- Koonin, E. V. (2015). Origin of eukaryotes from within archaea, archaeal eukaryome and bursts of gene gain: Eukaryogenesis just made easier? *Philos. Trans. R. Soc. B Biol. Sci.* 370:20140333. doi: 10.1098/rstb.2014.0333
- Koonin, E. V. (2016). Viruses and mobile elements as drivers of evolutionary transitions. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 371:20150442. doi: 10.1098/RSTB.2015.0442
- Koonin, E. V., and Krupovic, M. (2020). Phages build anti-defence barriers. *Nat. Microbiol.* 5, 8–9. doi: 10.1038/S41564-019-0635-Y
- Koonin, E. V., and Yutin, N. (2018). Multiple evolutionary origins of giant viruses. *F1000Res* 7. doi: 10.12688/F1000RESEARCH.16248.1
- Koonin, E. V., Dolja, V. V., and Krupovic, M. (2022). The logic of virus evolution. *Cell Host Microbe* 30, 917–929. doi: 10.1016/j.chom.2022.06.008
- Koonin, E. V., Dolja, V. V., Krupovic, M., and Kuhn, J. H. (2021). Viruses defined by the position of the virosphere within the replicator space. *Microbiol. Mol. Biol. Rev.* 85:e0019320. doi: 10.1128/MMBR.00193-20
- Koonin, E. V., Dolja, V. V., Krupovic, M., Varsani, A., Wolf, Y. I., Yutin, N., et al. (2020). Global organization and proposed megataxonomy of the virus world. *Microbiol. Mol. Biol. Rev.* 84, 1–33. doi: 10.1128/MMBR.00061-19
- Koonin, E. V., Senkevich, T. G., and Dolja, V. V. (2006). The ancient virus world and evolution of cells. *Biol. Direct* 1:29. doi: 10.1186/1745-6150-1-29
- Koonin, E. V., Wolf, Y. I., and Katsnelson, M. I. (2017). Inevitability of the emergence and persistence of genetic parasites caused by evolutionary instability of parasite-free states. *Biol. Direct* 12:31. doi: 10.1186/S13062-017-0202-5
- Koskella, B., and Brockhurst, M. A. (2014). Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol. Rev.* 38, 916–931. doi: 10.1111/1574-6976.12072
- Krakauer, D. C., and Sasaki, A. (2002). Noisy clues to the origin of life. *Proceedings Biol. Sci.* 269, 2423–2428. doi: 10.1098/RSPB.2002.2127
- Krishnamurthy, S. R., and Wang, D. (2017). Origins and challenges of viral dark matter. *Virus Res.* 239, 136–142. doi: 10.1016/j.virusres.2017.02.002
- Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982). Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell* 31, 147–157. doi: 10.1016/0092-8674(82)90414-7
- Krupovic, M., and Bamford, D. H. (2008). Virus evolution: How far does the double  $\beta$ -barrel viral lineage extend? *Nat. Rev. Microbiol.* 6, 941–948. doi: 10.1038/nrmicro2033
- Krupovic, M., and Koonin, E. V. (2017). Multiple origins of viral capsid proteins from cellular ancestors. *Proc. Natl. Acad. Sci. U.S.A.* 114, E2401–E2410. doi: 10.1073/PNAS.1621061114
- Krupovic, M., Dolja, V. V., and Koonin, E. V. (2019). Origin of viruses: Primordial replicators recruiting capsids from hosts. *Nat. Rev. Microbiol.* 17, 449–458. doi: 10.1038/S41579-019-0205-6
- Krupovic, M., Quemin, E. R. J., Bamford, D. H., Forterre, P., and Prangishvili, D. (2014). Unification of the globally distributed spindle-shaped viruses of the archaea. *J. Virol.* 88, 2354–2358. doi: 10.1128/jvi.02941-13
- La Scola, B., Audic, S., Robert, C., Jungang, L., De Lamballerie, X., Drancourt, M., et al. (2003). A giant virus in amoebae. *Science* 299:2033. doi: 10.1126/SCIENCE.1081867
- Laber, C. P., Hunter, J. E., Carvalho, F., Collins, J. R., Hunter, E. J., Schieler, B. M., et al. (2018). Coccolithovirus facilitation of carbon export in the North Atlantic. *Nat. Microbiol.* 3, 537–547. doi: 10.1038/s41564-018-0128-4
- Laidler, J. R., and Stedman, K. M. (2010). Virus silicification under simulated hot spring conditions. *Astrobiology* 10, 569–576. doi: 10.1089/ast.2010.0463
- Laidler, J. R., Shugart, J. A., Cady, S. L., Bahjat, K. S., and Stedman, K. M. (2013). Reversible inactivation and desiccation tolerance of silicified viruses. *J. Virol.* 87, 13927–13929. doi: 10.1128/jvi.02825-13
- Lázaro, E., Arribas, M., Cabanillas, L., Román, I., and Acosta, E. (2018). Evolutionary adaptation of an RNA bacteriophage to the simultaneous increase in the within-host and extracellular temperatures. *Sci. Rep.* 8:8080. doi: 10.1038/S41598-018-26443-Z
- Lee, K. Y., and Lee, B. J. (2017). Structural and biochemical properties of novel self-cleaving ribozymes. *Molecules* 22:678. doi: 10.3390/molecules22040678
- Levisohn, R., and Spiegelman, S. (1969). Further extracellular Darwinian experiments with replicating RNA molecules: Diverse variants isolated under different selective conditions. *Proc. Natl. Acad. Sci. U.S.A.* 63, 805–811. doi: 10.1073/PNAS.63.3.805
- Lindell, D., Jaffe, J. D., Johnson, Z. I., Church, G. M., and Chisholm, S. W. (2005). Photosynthesis genes in marine viruses yield proteins during host infection. *Nature* 438, 86–89. doi: 10.1038/nature04111
- Lindell, D., Sullivan, M. B., Johnson, Z. I., Tolonen, A. C., Rohwer, F., and Chisholm, S. W. (2004). Transfer of photosynthesis genes to and from *Prochlorococcus* viruses. *Proc. Natl. Acad. Sci. U.S.A.* 101, 11013–11018. doi: 10.1073/pnas.0401526101
- Lingam, M., and Loeb, A. (2021). *Life in the cosmos?: From biosignatures to technosignatures*. Cambridge, MA: Harvard University Press.
- Maggiore, C., Stromberg, J., Blanco, Y., Goordial, J., Cloutis, E., García-Villadangos, M., et al. (2020). The limits, capabilities, and potential for life detection with MinION sequencing in a paleochannel mars analog. *Astrobiology* 20, 375–393. doi: 10.1089/ast.2018.1964
- Mann, N. H., Cook, A., Millard, A., Bailey, S., and Clokie, M. (2003). Bacterial photosynthesis genes in a virus. *Nature* 424, 741–741. doi: 10.1038/424741a
- Marchesi, A., Umeda, K., Komekawa, T., Matsubara, T., Flechsig, H., Ando, T., et al. (2021). An ultra-wide scanner for large-area high-speed atomic force microscopy with megapixel resolution. *Sci. Rep.* 11:13003. doi: 10.1038/s41598-021-92365-y
- Márquez, L. M., Redman, R. S., Rodriguez, R. J., and Roossinck, M. J. (2007). A virus in a fungus in a plant: Three-way symbiosis required for thermal tolerance. *Science* 315, 513–515. doi: 10.1126/science.1136237
- Matsuura, K., Watanabe, K., Matsuzaki, T., Sakurai, K., and Kimizuka, N. (2010). Self-assembled synthetic viral capsids from a 24-mer viral peptide fragment. *Angew. Chem. Int. Ed. Engl.* 49, 9662–9665. doi: 10.1002/anie.201004606
- McDonald, M. J. (2019). Microbial experimental evolution—a proving ground for evolutionary theory and a tool for discovery. *EMBO Rep.* 20:e46992. doi: 10.15252/EMBR.201846992
- McGee, L. W., Aitchison, E. W., Caudle, S. B., Morrison, A. J., Zheng, L., Yang, W., et al. (2014). Payoffs, not tradeoffs, in the adaptation of a virus to ostensibly conflicting selective pressures. *PLoS Genet.* 10:e1004611. doi: 10.1371/JOURNAL.PGEN.1004611

- Meents, A., and Wiedorn, M. O. (2019). Virus structures by X-Ray free-electron lasers. *Annu. Rev. Virol.* 6, 161–176. doi: 10.1146/ANNUREV-VIROLOGY-092818-015724
- Mills, D. R., Peterson, R. L., and Spiegelman, S. (1967). An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proc. Natl. Acad. Sci. U.S.A.* 58, 217–224. doi: 10.1073/PNAS.58.1.217
- Minicka, J., Elena, S. F., Borodynko-Filas, N., Rubiś, B., and Hasiów-Jaroszewska, B. (2017). Strain-dependent mutational effects for Pepino mosaic virus in a natural host. *BMC Evol. Biol.* 17:67. doi: 10.1186/s12862-017-0920-4
- Mizuuchi, R., Furubayashi, T., and Ichihashi, N. (2022). Evolutionary transition from a single RNA replicator to a multiple replicator network. *Nat. Commun.* 13:1460. doi: 10.1038/s41467-022-29113-x
- Mochizuki, T., Krupovic, M., Pehau-Arnaudet, G., Sako, Y., Forterre, P., and Prangishvili, D. (2012). Archaeal virus with exceptional virion architecture and the largest single-stranded DNA genome. *Proc. Natl. Acad. Sci. U.S.A.* 109, 13386–13391. doi: 10.1073/pnas.1203668109
- Moreira, D., and López-García, P. (2009). Ten reasons to exclude viruses from the tree of life. *Nat. Rev. Microbiol.* 7, 306–311. doi: 10.1038/nrmicro2108
- Morowitz, H., and Smith, E. (2007). Energy flow and the organization of life. *Complexity* 13, 51–59. doi: 10.1002/cplx.20191
- Mougari, S., Sahmi-Bounsiar, D., Levasseur, A., Colson, P., Scola, B., and La. (2019). Virophages of Giant viruses: An update at eleven. *Viruses* 11:733. doi: 10.3390/V11080733
- Muniesa, M., Imamovic, L., and Jofre, J. (2011). Bacteriophages and genetic mobilization in sewage and faecally polluted environments. *Microb. Biotechnol.* 4, 725–734. doi: 10.1111/j.1751-7915.2011.00264.x
- Munson-McGee, J. H., Peng, S., Dewerff, S., Stepanauskas, R., Whitaker, R. J., Weitz, J. S., et al. (2018). A virus or more in (nearly) every cell: Ubiquitous networks of virus-host interactions in extreme environments. *ISME J.* 12, 1706–1714. doi: 10.1038/s41396-018-0071-7
- Munson-McGee, J. H., Rooney, C., and Young, M. J. (2020). An uncultivated virus infecting a nanoarchaeal parasite in the hot springs of Yellowstone National Park. *J. Virol.* 94, e1213–e1219. doi: 10.1128/jvi.01213-19
- Mushegian, A. R. (2020). Are there 10<sup>31</sup> virus particles on earth, or more, or fewer? *J. Bacteriol.* 202, e52–e20. doi: 10.1128/JB.00052-20
- Navas-Castillo, J. (2009). Six comments on the ten reasons for the demotion of viruses. *Nat. Rev. Microbiol.* 7:615. doi: 10.1038/NRMICRO2108-C2
- Neri, U., Wolf, Y. I., Roux, S., Camargo, A. P., Lee, B., Kazlauskas, D., et al. (2022). Expansion of the global RNA virome reveals diverse clades of bacteriophages. *Cell* 185, 4023–4037. doi: 10.1016/j.cell.2022.08.023
- Nielsen, P. E. (2007). Peptide nucleic acids and the origin of life. *Chem. Biodivers.* 4, 1996–2002. doi: 10.1002/CBDV.200790166
- Nishida, R., Nakamura, K., Taniguchi, I., Murase, K., Ooka, T., Ogura, Y., et al. (2021). The global population structure and evolutionary history of the acquisition of major virulence factor-encoding genetic elements in Shiga toxin-producing *Escherichia coli* O121:H19. *Microb. Genomics* 7:000716. doi: 10.1099/MGEN.0.000716
- O’rourke, A., Zoumplis, A., Wilburn, P., Lee, M. D., Lee, Z., Vecina, M., et al. (2020). Following the astrobiology roadmap: Origins, habitability and future exploration. *Curr. Issues Mol. Biol.* 38, 1–32. doi: 10.21775/cimb.038.001
- Orange, F., Chabin, A., Gorlas, A., Lucas-Staat, S., Geslin, C., Le Romancer, M., et al. (2011). Experimental fossilisation of viruses from extremophilic Archaea. *Biogeosciences* 8, 1465–1475. doi: 10.5194/bg-8-1465-2011
- Oro, J., Miller, S. L., and Lazcano, A. (1990). The origin and early evolution of life on earth. *Annu. Rev. Earth Planet. Sci.* 18, 317–356. doi: 10.1146/ANNUREV.EA.18.050190.001533
- Pandey, S., Bean, R., Sato, T., Poudyal, I., Bielecki, J., Cruz Villarreal, J., et al. (2020). Time-resolved serial femtosecond crystallography at the European XFEL. *Nat. Methods* 17, 73–78. doi: 10.1038/s41592-019-0628-z
- Penadés, J. R., Chen, J., Quiles-Puchalt, N., Carpena, N., and Novick, R. P. (2015). Bacteriophage-mediated spread of bacterial virulence genes. *Curr. Opin. Microbiol.* 23, 171–178. doi: 10.1016/j.mib.2014.11.019
- Pérez-Losada, M., Arenas, M., Galán, J. C., Palero, F., and González-Candelas, F. (2015). Recombination in viruses: Mechanisms, methods of study, and evolutionary consequences. *Infect. Genet. Evol.* 30, 296–307. doi: 10.1016/j.meegid.2014.12.022
- Pfeiffer, J. K., and Kirkegaard, K. (2003). A single mutation in poliovirus RNA-dependent RNA polymerase confers resistance to mutagenic nucleotide analogs via increased fidelity. *Proc. Natl. Acad. Sci. U.S.A.* 100, 7289–7294. doi: 10.1073/PNAS.1232294100
- Pietilä, M. K., Atanasova, N. S., Oksanen, H. M., and Bamford, D. H. (2013). Modified coat protein forms the flexible spindle-shaped virion of haloarchaeal virus His1. *Environ. Microbiol.* 15, 1674–1686. doi: 10.1111/1462-2920.12030
- Plunkett, G., Rose, D. J., Durfee, T. J., and Blattner, F. R. (1999). Sequence of Shiga toxin 2 phage 933W from *Escherichia coli* O157:H7: Shiga toxin as a phage late-gene product. *J. Bacteriol.* 181, 1767–1778. doi: 10.1128/JB.181.6.1767-1778.1999
- Poorvin, L., Rinta-Kanto, J. M., Hutchins, D. A., and Wilhelm, S. W. (2004). Viral release of iron and its bioavailability to marine plankton. *Limnol. Oceanogr.* 49, 1734–1741. doi: 10.4319/lo.2004.49.5.1734
- Quakkelaar, E. D., Beaumont, T., van Nuenen, A. C., van Alphen, F. P. J., Boeser-Nunnink, B. D. M., van ’t Wout, A. B., et al. (2007). T cell line passage can select for pre-existing neutralization-sensitive variants from the quasispecies of primary human immunodeficiency virus type-1 isolates. *Virology* 359, 92–104. doi: 10.1016/J.VIROL.2006.09.021
- Quemin, E. R. J., Pietilä, M. K., Oksanen, H. M., Forterre, P., Rijpstra, W. I. C., Schouten, S., et al. (2015). Sulfolobus spindle-shaped virus 1 contains glycosylated capsid proteins, a cellular chromatin protein, and host-derived lipids. *J. Virol.* 89, 11681–11691. doi: 10.1128/jvi.02270-15
- Rawn, S. M., and Cross, J. C. (2008). The evolution, regulation, and function of placenta-specific genes. *Annu. Rev. Cell Dev. Biol.* 24, 159–181. doi: 10.1146/ANNUREV.CELLBIO.24.110707.175418
- Reche, I., D’Orta, G., Mladenov, N., Winget, D. M., and Suttle, C. A. (2018). Deposition rates of viruses and bacteria above the atmospheric boundary layer. *ISME J.* 12, 1154–1162. doi: 10.1038/S41396-017-0042-4
- Rezzonico, F. (2014). Nanopore-based instruments as biosensors for future planetary missions. *Astrobiology* 14, 344–351. doi: 10.1089/ast.2013.1120
- Rice, G., Tang, L., Stedman, K., Roberto, F., Spuhler, J., Gillitzer, E., et al. (2004). The structure of a thermophilic archaeal virus shows a double-stranded DNA viral capsid type that spans all domains of life. *Proc. Natl. Acad. Sci. U.S.A.* 101, 7716–7720. doi: 10.1073/PNAS.0401773101
- Rodrigues, R. A. L., da Silva, L. C. F., and Abrahão, J. S. (2020). Translating the language of giants: translation-related genes as a major contribution of giant viruses to the virosphere. *Arch. Virol.* 165, 1267–1278. doi: 10.1007/s00705-020-04626-2
- Rosenzweig, J. B., Majernik, N., Robles, R. R., Andonian, G., Camacho, O., Fukasawa, A., et al. (2020). An ultra-compact x-ray free-electron laser. *N. J. Phys.* 22:093067. doi: 10.1088/1367-2630/abb16c
- Roux, S., Brum, J. R., Dutilh, B. E., Sunagawa, S., Duhaime, M. B., Loy, A., et al. (2016). Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature* 537, 689–693. doi: 10.1038/NATURE19366
- Rowe, H. M., and Trono, D. (2011). Dynamic control of endogenous retroviruses during development. *Virology* 411, 273–287. doi: 10.1016/J.VIROL.2010.12.007
- Ruiz-Jarabo, C. M., Arias, A., Molina-Paris, C., Briones, C., Baranowski, E., Escarmis, C., et al. (2002). Duration and fitness dependence of quasispecies memory. *J. Mol. Biol.* 315, 285–296. doi: 10.1006/jmbi.2001.5232
- Russell, M. J., and Hall, A. J. (1997). The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and pH front. *J. Geol. Soc. Lond.* 154, 377–402. doi: 10.1144/GSJS.154.3.0377
- Sanjuán, R., Moya, A., and Elena, S. F. (2004). The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8396–8401. doi: 10.1073/PNAS.0400146101
- Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M., and Belshaw, R. (2010). Viral mutation rates. *J. Virol.* 84, 9733–9748. doi: 10.1128/JVI.00694-10
- Sankaran, N. (2016). The RNA world at thirty: A look back with its author. *J. Mol. Evol.* 83, 169–175. doi: 10.1007/s00239-016-9767-3
- Santos, F., Yarza, P., Parro, V., Meseguer, I., Rosselló-Móra, R., and Antón, J. (2012). Culture-independent approaches for studying viruses from hypersaline environments. *Appl. Environ. Microbiol.* 78, 1635–1643. doi: 10.1128/AEM.07175-11
- Santos-Pérez, I., Charro, D., Gil-Carton, D., Azkargorta, M., Elortza, F., Bamford, D. H., et al. (2019). Structural basis for assembly of vertical single β-barrel viruses. *Nat. Commun.* 10:1184. doi: 10.1038/S41467-019-08927-2
- Sawström, C., Lisle, J., Anesio, A. M., Priscu, J. C., and Laybourn-Parry, J. (2008). Bacteriophage in polar inland waters. *Extremophiles* 12, 167–175. doi: 10.1007/s00792-007-0134-6
- Schulz, F., Yutin, N., Ivanova, N. N., Ortega, D. R., Lee, T. K., Vierheilig, J., et al. (2017). Giant viruses with an expanded complement of translation system components. *Science* 356, 82–85. doi: 10.1126/SCIENCE.AAL4657

- Seibert, M. M., Ekeberg, T., Maia, F. R. N. C., Svenda, M., Andreasson, J., Jónsson, O., et al. (2011). Single mimivirus particles intercepted and imaged with an X-ray laser. *Nature* 470, 78–82. doi: 10.1038/nature09748
- Sharma, G., and Curtis, P. D. (2022). The impacts of microgravity on bacterial metabolism. *Life (Basel)* 12:774. doi: 10.3390/LIFE12060774
- Sharon, I., Tzahor, S., Williamson, S., Shmoish, M., Man-Aharonovich, D., Rusch, D. B., et al. (2007). Viral photosynthetic reaction center genes and transcripts in the marine environment. *ISME J.* 1, 492–501. doi: 10.1038/ISMEJ.2007.67
- Shelford, E. J., Middelboe, M., Møller, E. F., and Suttle, C. A. (2012). Virus-driven nitrogen cycling enhances phytoplankton growth. *Aquat. Microb. Ecol.* 66, 41–46. doi: 10.3354/ame01553
- Shi, M., Lin, X. D., Tian, J. H., Chen, L. J., Chen, X., Li, C. X., et al. (2016). Redefining the invertebrate RNA virosphere. *Nature* 540, 539–543. doi: 10.1038/nature20167
- Shoemaker, S. C., and Ando, N. (2018). X-rays in the Cryo-EM era: Structural biology's dynamic future. *Biochemistry* 57:277. doi: 10.1021/ACS.BIOCHEM.7B01031
- Smith, P. H., Tamppari, L., Arvidson, R. E., Bass, D., Blaney, D., Boynton, W., et al. (2008). Introduction to special section on the phoenix mission: Landing site characterization experiments, mission overviews, and expected science. *J. Geophys. Res.* 113:E00A18. doi: 10.1029/2008JE003083
- Somovilla, P., Rodríguez-Moreno, A., Arribas, M., Manrubia, S., and Lázaro, E. (2022). Standing genetic diversity and transmission bottleneck size drive adaptation in bacteriophage Q $\beta$ . *Int. J. Mol. Sci.* 23:8876. doi: 10.3390/IJMS23168876
- Stanley, S. Y., and Maxwell, K. L. (2018). Phage-encoded anti-CRISPR defenses. *Annu. Rev. Genet.* 52, 445–464. doi: 10.1146/annurev-genet-120417-031321
- Starr, E. P., Nuccio, E. E., Pett-Ridge, J., Banfield, J. F., and Firestone, M. K. (2019). Metatranscriptomic reconstruction reveals RNA viruses with the potential to shape carbon cycling in soil. *Proc. Natl. Acad. Sci. U.S.A.* 116, 25900–25908. doi: 10.1073/pnas.1908291116
- Stedman, K. M., Clore, A. J., and Combet-Blanc, Y. (2006). “Biogeographical diversity of archaeal viruses,” in *GM symposium 66: Prokaryotic diversity: Mechanisms and significance*, eds N. A. Logan, H. M. Lappin-Scott, and P. C. F. Oyston (Cambridge: Cambridge University Press), 131–144.
- Sullivan, M. B., Coleman, M. L., Weigle, P., Rohwer, F., and Chisholm, S. W. (2005). Three *Prochlorococcus* cyanophage genomes: Signature features and ecological interpretations. *PLoS Biol.* 3:e144. doi: 10.1371/journal.pbio.0030144
- Sullivan, M. B., Weitz, J. S., and Wilhelm, S. (2017). Viral ecology comes of age. *Environ. Microbiol. Rep.* 9, 33–35. doi: 10.1111/1758-2229.12504
- Sunagawa, S., Acinas, S. G., Bork, P., Bowler, C., Acinas, S. G., Babin, M., et al. (2020). Tara oceans: Towards global ocean ecosystems biology. *Nat. Rev. Microbiol.* 18, 428–445. doi: 10.1038/s41579-020-0364-5
- Suttle, C. A. (1994). The significance of viruses to mortality in aquatic microbial communities. *Microb. Ecol.* 28, 237–243. doi: 10.1007/BF00166813
- Suttle, C. A. (2005). Viruses in the sea. *Nature* 437, 356–361. doi: 10.1038/nature04160
- Suttle, C. A. (2007). Marine viruses—major players in the global ecosystem. *Nat. Rev. Microbiol.* 5, 801–812. doi: 10.1038/nrmicro1750
- Suzuki, Y., Hino, H., Hawai, T., Saito, K., Kotsugi, M., and Ono, K. (2020). Symmetry prediction and knowledge discovery from X-ray diffraction patterns using an interpretable machine learning approach. *Sci. Rep.* 10, 1–11. doi: 10.1038/s41598-020-77474-4
- Swanson, M. M., Fraser, G., Daniell, T. J., Torrance, L., Gregory, P. J., and Taliany, M. (2009). Viruses in soils: Morphological diversity and abundance in the rhizosphere. *Ann. Appl. Biol.* 155, 51–60. doi: 10.1111/j.1744-7348.2009.00319.x
- Szathmáry, E. (2015). Toward major evolutionary transitions theory 2.0. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10104–10111. doi: 10.1073/PNAS.1421398112
- Szathmáry, E., and Demeter, L. (1987). Group selection of early replicators and the origin of life. *J. Theor. Biol.* 128, 463–486. doi: 10.1016/S0022-5193(87)80191-1
- Szathmáry, E., and Smith, J. M. (1995). The major evolutionary transitions. *Nature* 374, 227–232. doi: 10.1038/374227A0
- Szostak, J. W. (2012). Attempts to define life do not help to understand the origin of life. *J. Biomol. Struct. Dyn.* 29, 599–600. doi: 10.1080/073911012010524998
- Szostak, J. W. (2017). The narrow road to the deep past: In search of the chemistry of the origin of life. *Angew. Chem. Int. Ed. Engl.* 56, 11037–11043. doi: 10.1002/ANIE.201704048
- Takemura, M. (2020). Medusavirus ancestor in a proto-eukaryotic cell: Updating the hypothesis for the viral origin of the nucleus. *Front. Microbiol.* 11:571831. doi: 10.3389/FMICB.2020.571831
- Thierauf, A., Perez, G., and Maloy, A. S. (2009). Generalized transduction. *Methods Mol. Biol.* 501, 267–286. doi: 10.1007/978-1-60327-164-6\_23
- Thingstad, T. F., and Lignell, R. (1997). Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. *Aquat. Microb. Ecol.* 13, 19–27. doi: 10.3354/ame013019
- Thomas, E., Anderson, R. E., Li, V., Rogan, L. J., and Huber, J. A. (2021). Diverse viruses in deep-sea hydrothermal vent fluids have restricted dispersal across ocean basins. *mSystems* 6, e68–e21. doi: 10.1128/MSYSTEMS.00068-21
- Ting, C. N., Rosenberg, M. P., Snow, C. M., Samuelson, L. C., and Meisler, M. H. (1992). Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene. *Genes Dev.* 6, 1457–1465. doi: 10.1101/GAD.6.8.1457
- Trubl, G., Jang, H., Bin, Roux, S., Emerson, J. B., Solonenko, N., Vik, D. R., et al. (2018). Soil viruses are underexplored players in ecosystem carbon processing. *mSystems* 3, e76–e18. doi: 10.1128/msystems.00076-18
- Trus, B. L., Cheng, N., Newcomb, W. W., Homa, F. L., Brown, J. C., and Steven, A. C. (2004). Structure and polymorphism of the UL6 portal protein of herpes simplex virus type 1. *J. Virol.* 78, 12668–12671. doi: 10.1128/JVI.78.22.12668-12671.2004
- Vaishampayan, A., and Grohmann, E. (2019). Multi-resistant biofilm-forming pathogens on the international space station. *J. Biosci.* 44:125. doi: 10.1007/s12038-019-9929-8
- Vale, P. F., Choisy, M., Froissart, R., Sanjuán, R., and Gandon, S. (2012). The distribution of mutational fitness effects of phage  $\phi$ X174 on different hosts. *Evolution* 66, 3495–3507. doi: 10.1111/J.1558-5646.2012.01691.X
- Van den Bergh, B., Swings, T., Fauvar, M., and Michiels, J. (2018). Experimental design, population dynamics, and diversity in microbial experimental evolution. *Microbiol. Mol. Biol. Rev.* 82:e00008–e18. doi: 10.1128/mmbrev.00008-18
- Van Etten, J. L., and Meints, R. H. (1999). Giant viruses infecting algae. *Annu. Rev. Microbiol.* 53, 447–494. doi: 10.1146/ANNUREV.MICRO.53.1.447
- Vignuzzi, M., Stone, J. K., Arnold, J. J., Cameron, C. E., and Andino, R. (2006). Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* 439, 344–348. doi: 10.1038/NATURE04388
- Villarreal, L. P., and Witzany, G. (2010). Viruses are essential agents within the roots and stem of the tree of life. *J. Theor. Biol.* 262, 698–710. doi: 10.1016/J.JTBI.2009.10.014
- Vitas, M., and Dobovišek, A. (2019). Towards a general definition of life. *Orig. Life Evol. Biosph.* 49, 77–88. doi: 10.1007/s11084-019-09578-5
- Waddell, T. E., Franklin, K., Mazzocco, A., Kropinski, A. M., and Johnson, R. P. (2009). Generalized transduction by lytic bacteriophages. *Methods Mol. Biol.* 501, 293–303. doi: 10.1007/978-1-60327-164-6\_25
- Wang, W., Zhao, H., and Han, G.-Z. (2020b). Host-virus arms races drive elevated adaptive evolution in viral receptors. *J. Virol.* 94, e684–e620. doi: 10.1128/jvi.00684-20
- Wang, F., Baquero, D. P., Beltran, L. C., Su, Z., Osinski, T., Zheng, W., et al. (2020a). Structures of filamentous viruses infecting hyperthermophilic archaea explain DNA stabilization in extreme environments. *Proc. Natl. Acad. Sci. U.S.A.* 117, 19643–19652. doi: 10.1073/PNAS.2011125117
- Wasik, B. R., Bhushan, A., Ogbunugafor, C. B., and Turner, P. E. (2015). Delayed transmission selects for increased survival of vesicular stomatitis virus. *Evolution* 69, 117–125. doi: 10.1111/EVO.12544
- Watson, B. N. J., Steens, J. A., Staals, R. H. J., Westra, E. R., and van Houte, S. (2021). Coevolution between bacterial CRISPR-Cas systems and their bacteriophages. *Cell Host Microbe* 29, 715–725. doi: 10.1016/j.chom.2021.03.018
- Watson, J. D., and Crick, F. H. C. (1956). Structure of small viruses. *Nature* 177, 473–475.
- Weiss, M. C., Sousa, F. L., Mrnjavac, N., Neukirchen, S., Roettger, M., Nelson-Sathi, S., et al. (2016). The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* 1, 1–8. doi: 10.1038/nmicrobiol.2016.116
- Weiss, R. A. (2017). Exchange of genetic sequences between viruses and hosts. *Curr. Top. Microbiol. Immunol.* 407, 1–29. doi: 10.1007/82\_2017\_21
- Whitaker, R. J., Grogan, D. W., and Taylor, J. W. (2003). Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science* 301, 976–978. doi: 10.1126/science.1086909
- White, R. A., Visscher, P. T., and Burns, B. P. (2021). Between a rock and a soft place: The role of viruses in lithification of modern microbial mats. *Trends Microbiol.* 29, 204–213. doi: 10.1016/J.TIM.2020.06.004

- Whittington, A. C., and Rokyta, D. R. (2019). Biophysical spandrels form a hot-spot for kosmotropic mutations in bacteriophage thermal adaptation. *J. Mol. Evol.* 87, 27–36. doi: 10.1007/S00239-018-9882-4
- Wiedenheft, B., Stedman, K., Roberto, F., Willits, D., Gleske, A.-K., Zoeller, L., et al. (2004). Comparative genomic analysis of hyperthermophilic archaeal fuselloviridae viruses. *J. Virol.* 78, 1954–1961. doi: 10.1128/jvi.78.4.1954-1961.2004
- Wigington, C. H., Sonderegger, D., Brussaard, C. P. D., Buchan, A., Finke, J. F., Fuhrman, J. A., et al. (2016). Re-examination of the relationship between marine virus and microbial cell abundances. *Nat. Microbiol.* 1, 4–11. doi: 10.1038/nmicrobiol.2015.24
- Wilhelm, S. W., and Suttle, C. A. (1999). Viruses and nutrient cycles in the sea. *Bioscience* 49, 781–788. doi: 10.2307/1313569
- Woese, C. R. (1987). Bacterial evolution. *Microbiol. Rev.* 51, 221–271. doi: 10.1128/MR.51.2.221-271.1987
- Woese, C. R. (2002). On the evolution of cells. *Proc. Natl. Acad. Sci. U.S.A.* 99, 8742–8747. doi: 10.1073/PNAS.132266999
- Woese, C. R., Kandler, O., and Wheelis, M. L. (1990). Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. U.S.A.* 87, 4576–4579. doi: 10.1073/PNAS.87.12.4576
- Wolf, D., and Goff, S. P. (2008). Host restriction factors blocking retroviral replication. *Annu. Rev. Genet.* 42, 143–163. doi: 10.1146/ANNUREV.GENET.42.110807.091704
- Wolf, Y. I., Silas, S., Wang, Y., Wu, S., Bocek, M., Kazlauskas, D., et al. (2020). Doubling of the known set of RNA viruses by metagenomic analysis of an aquatic virome. *Nat. Microbiol.* 5, 1262–1270. doi: 10.1038/s41564-020-0755-4
- Yang, B., Fang, L., Gao, Q., Xu, C., Xu, J., Chen, Z. X., et al. (2022). Species-specific KRAB-ZFPs function as repressors of retroviruses by targeting PBS regions. *Proc. Natl. Acad. Sci. U.S.A.* 119:e2119415119. doi: 10.1073/pnas.2119415119
- Yooshep, S., Neelson, K. H., Rusch, D. B., McCrow, J. P., Dupont, C. L., Kim, M., et al. (2010). Genomic and functional adaptation in surface ocean planktonic prokaryotes. *Nature* 468, 60–66. doi: 10.1038/nature09530
- Yoshida-Takashima, Y., Nunoura, T., Kazama, H., Noguchi, T., Inoue, K., Akashi, H., et al. (2012). Spatial distribution of viruses associated with planktonic and attached microbial communities in hydrothermal environments. *Appl. Environ. Microbiol.* 78, 1311–1320. doi: 10.1128/AEM.06491-11
- Yutin, N., Bäckström, D., Ettema, T. J. G., Krupovic, M., and Koonin, E. V. (2018). Vast diversity of prokaryotic virus genomes encoding double jelly-roll major capsid proteins uncovered by genomic and metagenomic sequence analysis. *Virol. J.* 15:67. doi: 10.1186/S12985-018-0974-Y
- Zablocki, O., Adriaenssens, E. M., and Cowan, D. (2015). Diversity and ecology of viruses in hyperarid desert soils. *Appl. Environ. Microbiol.* 82, 770–777. doi: 10.1128/AEM.02651-15
- Zandi, R., Reguera, D., Bruinsma, R. F., Gelbart, W. M., and Rudnick, J. (2004). Origin of icosahedral symmetry in viruses. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15556–15560. doi: 10.1073/pnas.0405844101
- Zayed, A. A., Wainaina, J. M., Dominguez-Huerta, G., Pelletier, E., Guo, J., Mohssen, M., et al. (2022). Cryptic and abundant marine viruses at the evolutionary origins of earth's RNA virome. *Science* 376, 156–162. doi: 10.1126/science.abm5847
- Zhang, Y. Z., Shi, M., and Holmes, E. C. (2018). Using metagenomics to characterize an expanding virosphere. *Cell* 172, 1168–1172. doi: 10.1016/j.cell.2018.02.043
- Zhao, L., Seth-Pasricha, M., Stemate, D., Crespo-Bellido, A., Gagnon, J., Draghi, J., et al. (2019). Existing host range mutations constrain further emergence of RNA viruses. *J. Virol.* 93, e1385–e1318. doi: 10.1128/JVI.01385-18
- Zhao, Y., Temperton, B., Thrash, J. C., Schwabach, M. S., Vergin, K. L., Landry, Z. C., et al. (2013). Abundant SAR11 viruses in the ocean. *Nature* 494, 357–360. doi: 10.1038/nature11921
- Zheng, L., Liang, X., Shi, R., Li, P., Zhao, J., Li, G., et al. (2020). Viral abundance and diversity of production fluids in oil reservoirs. *Microorganisms* 8, 1–15. doi: 10.3390/MICROORGANISMS8091429
- Zimmerman, A. E., Howard-Varona, C., Needham, D. M., John, S. G., Worden, A. Z., Sullivan, M. B., et al. (2020). Metabolic and biogeochemical consequences of viral infection in aquatic ecosystems. *Nat. Rev. Microbiol.* 18, 21–34. doi: 10.1038/s41579-019-0270-x