Astrovirology: Viruses at Large in the Universe.

Permalink
https://escholarship.org/uc/item/4d87k128

Journal
Astrobiology, 18(2)

ISSN
1531-1074

Authors
Berliner, Aaron J
Mochizuki, Tomohiro
Stedman, Kenneth M

Publication Date
2018-02-01

DOI
10.1089/ast.2017.1649

Peer reviewed
Astrovirology: Viruses at Large in the Universe

Aaron J. Berliner,1 Tomohiro Mochizuki,2 and Kenneth M. Stedman3

Abstract

Viruses are the most abundant biological entities on modern Earth. They are highly diverse both in structure and genomic sequence, play critical roles in evolution, strongly influence terran biogeochemistry, and are believed to have played important roles in the origin and evolution of life. However, there is yet very little focus on viruses in astrobiology. Viruses arguably have coexisted with cellular life-forms since the earliest stages of life, may have been directly involved therein, and have profoundly influenced cellular evolution. Viruses are the only entities on modern Earth to use either RNA or DNA in both single- and double-stranded forms for their genetic material and thus may provide a model for the putative RNA-protein world. With this review, we hope to inspire integration of virus research into astrobiology and also point out pressing unanswered questions in astrovirology, particularly regarding the detection of virus biosignatures and whether viruses could be spread extraterrestrially. We present basic virology principles, an inclusive definition of viruses, review current virology research pertinent to astrobiology, and propose ideas for future astrovirology research foci. Key Words: Astrobiology—Virology—Biosignatures—Origin of life—Roadmap. Astrobiology 18, 207–223.

1.1. Where Are Viruses in Astrobiology?

We consider viruses to be central to multiple aspects of astrobiology, but in the current NASA Astrobiology Strategy (Hays et al., 2015), viruses are mentioned only six times in over 250 pages. The European AstRoMap is slightly more inclusive, with nine mentions of viral elements (two of them parenthetical) in 42 pages but none of viruses (Horneck et al., 2016). Perhaps the astrobiology community is not yet fully aware of the amazing ubiquity, diversity, and roles of viruses in biogeochemistry, evolution, and the origins of life. Therefore, we provide an annotated version of the NASA Astrobiology Strategy Goals with a focus on viruses and emphasize their critical and underappreciated roles in astrobiology. By way of background we also provide an overview of viruses and their replication, while focusing on aspects most relevant to astrobiology.

1.2. An Astrovirology Strategy

The 2015 NASA Astrobiology Strategy provides a list of goals for the astrobiology community which aim to improve research products on the origins, evolution, and future of life on Earth and beyond, while also aiming to develop stronger communication across disciplinary boundaries. Table 1 describes these goals and subgoals in relationship to how virology can be considered an integral component of astrobiology.

1.3. The Central Role of Viruses in Astrobiology

The extracellular form of viruses, virions, are the most abundant biological particles on Earth, with an estimated 10^31 virions in the oceans alone (Suttle, 2005). Whether this amazing virion abundance is true in extraterrestrial oceans or was true in primordial Earth oceans are open questions. Unfortunately, the lack of validated virus biosignatures makes detection of extraterrestrial or ancient viruses challenging. Virus infections profoundly influence both global and local biogeochemical cycling (Suttle, 2009) yet have not been included in ancient climate models or models for extraterrestrial climates. Viruses are critical drivers of current evolution (Enard et al., 2016) and were probably extremely important for the early evolution of life on Earth (Koonin et al., 2006). These roles of viruses in astrobiology are described in more detail in Section 3. For readers who are less familiar with viruses and virus replication cycles, we describe the history of astrovirology, virus replication, virion structures, and genomes in Section 2.

1University of California, Berkeley, California, USA.
2Earth Life Sciences Institute, Tokyo Institute of Technology, Japan.
3Center for Life in Extreme Environments and Biology Department, Portland State University, Oregon, USA.
Table 1. **Astrovirology Goals in the Framework of the 2015 NASA Astrobiology Strategy**

<table>
<thead>
<tr>
<th>Original goal</th>
<th>Subgoals</th>
<th>Astrovirology subgoals</th>
</tr>
</thead>
</table>
| Foster Interdisciplinary Science | - Understand linkages between geology, climate, and life.  
- Coordinate field/mission observations, laboratory experimentation, and theory.  
- Enhance effective communication between scientists and technologists.  
- Promote collaborations between government offices and agencies.  
- Enhance the ability of society to address important challenges that require interdisciplinary approaches. | - Understand the linkages between virology, geology, climate, and life.  
- Coordinate field/mission observations, laboratory experimentation, and theory.  
- Enhance communication with virologists.  
- Promote collaborations with virologists.  
- Enhance societal understanding of the roles of viruses on our planet. |
| Enhance NASA’s Missions | - Research analog environments to understand processes underlying habitability.  
- Develop more effective proxies for life (biosignatures) and environments (“envirosignatures”). | - Develop broader proxies for viruses as indirect evidence of life in analog environments.  
- Develop and validate virus biosignatures. |
| Promote Planetary Stewardship | - Understand interactions between biological activity and environmental change at several spatial scales.  
- Support planetary protection efforts.  
- Constrain human exploration (minimize adverse impacts on other planetary environments). | - Understand virus-environment interactions at multiple spatial scales.  
- Evaluate the roles of viruses in planetary protection. |
| Enhance Societal Interest and Relevance | - Act upon the public interest to understand our place in the Universe and broaden understanding of life.  
- Sustain the human imperative to explore the unknown. | - Increase understanding of our place in the Universe through understanding our relationships with viruses and their past and present effects on evolution and life.  
- Explore the mostly unknown world (or worlds) of viruses.  
- Increase interest in science and technology by including (astro)virologists.  
- Increase benefits to society through application of astrovirology technologies.  
- Incorporate virology perspectives to enhance critical and creative thinking. |
| Inspire Future Generations | - Enhance the interest in science and technology to the benefit of our society.  
- Enhance critical and creative thinking in all fields of inquiry. |  
  
### 1.4. The History of “Astrovirology”

The first to publicly connect the study of viruses and astrobiology was the late Baruch Blumberg, discoverer of the hepatitis B virus, Nobel laureate, and first director of the NASA Astrobiology Institute (NAI), who made the connection at the Astrobiology Science Conference in 2002. Thereafter, NAI formed the Virus Focus Group, NAIViFoG. NAIViFoG has organized a number of workshops and symposia mostly focusing on virus origins, ecology, and evolution (Stedman and Blumberg, 2005). The first use of the term “astrovirology” to our knowledge was as a session and abstract title at AbSciCon 2004 (Blumberg, 2004). Dale W. Griffin used “Astrovirology” as a subheading in his 2013 essay in *Astrobiology*: “The Quest for Extraterrestrial Life: What about the Viruses?” We completely agree with his statement: “We should be looking for viruses in our quest for extraterrestrial life” (Griffin, 2013), but, to our knowledge, very few astrobiologists have even started looking.

### 2.1. What Is a Virus?

Before we start “looking for viruses in our quest for extraterrestrial life” (Griffin, 2013), however, we need to agree on what a virus actually is. The original virus definition was practical: a disease-causing particle that passed through a filter that retains all known bacteria “contagium vivum fluidum” (contagious living fluid) (Beijerinck, 1898). The standard definition of viruses for many decades thereafter was “a very small disease-causing agent.” Later, there were debates whether they were infectious liquid or particles, but the development of the electron microscope in the 1930s revealed that all viruses were found to have an inert extracellular particle state, the virion. Virions consist of a protein shell (also known as a capsid) that packages the viral genomic nucleic acid, which can be composed of DNA or RNA. In addition, some viruses can have a lipid-containing outer layer, called an envelope. Sir Peter Medawar, Nobel laureate and pioneering immunologist, is said to have called viruses “a piece of bad news wrapped up in a protein.” All viruses must transfer their genome inside a host cell and reprogram that cell to produce more virus. This discovery led to the expansion of the standard virus definition to “a very small obligate intracellular parasite” (Acheson, 2011). However, as discussed below, other than “obligate intracellular” this definition is incomplete. A more inclusive definition of viruses, that we prefer, was proposed by Salvador Luria and coauthors: “Viruses are entities whose...
2.3. Virus Replication Cycle

Viruses are elements of nucleic acid that replicate inside living cells using the cellular synthetic machinery and causing the synthesis of specialized elements that can transfer the viral genome to other cells” (Luria et al., 1978).

2.2. Viruses Are Not Just Virions: Are Viruses Alive?

It is important to make the distinction between a virus and a virion. The inert virion, the form visible under the electron microscope, is analogous to a seed or a spore that can only replicate in an appropriate environment, in the case of a virus inside a host cell. Virus refers to the whole virus replication cycle (see Luria’s definition above). Unfortunately, most people, including many virologists, conflate the two. Considering the virion to be a virus is analogous to considering an acorn an oak tree or a spermatozoid a human (Claverie, 2006; Forterre, 2010). When one considers the whole virus replication cycle (Fig. 1A), it comes close to NASA’s working definition of life: “A self-sustaining chemical system capable of Darwinian evolution.” Viruses are chemical, they are capable of Darwinian evolution, but they are not “self-sustaining” per se, as they require infection of a living cell for their replication. For further discussion of the vital nature of viruses, see the works of Schulz et al. (2017) and Koonin and Dolja (2014). Whether viruses are alive or not may be a moot question, but if a virion (or a virus-like particle) were to be unequivocally detected in an extraterrestrial sample, very few people would claim that this would not be evidence for life—wherever that sample was from.

2.3. Virus Replication Cycle

Following Luria’s description, a basic virus replication cycle consists of (1) binding of a virion to a host cell, (2) delivery of the viral genome into the host cell, (3) reprogramming of the cellular machinery to produce more virions, and (4) release of virions from the host cell to infect other cells. Then the process repeats. As a result, while cellular organisms generally replicate by binary fission, with other cells. Then the process repeats. As a result, while cellular organisms generally replicate by binary fission, the envelope, outside the protein capsid. Virus capsids generally have very simple elegant geometries, often helical symmetry, similar to a soccer ball (e.g., rhinovirus, known to cause the common cold), or helical symmetry (e.g., tobacco mosaic virus). Another common virus structure is the pleomorphic structure, often found with enveloped viruses (e.g., influenza virus) (Fig. 2). In addition, many virions have additional proteins on their surface that are used to recognize and bind to the host cell surface. These proteins can often be found at the vertices of polyhedral virions, either one or both ends of helical virions, or on the surface of enveloped virions.

Polyhedral and helical structures are widespread in the viral world, primarily for geometrical reasons. Both of these structures can be formed with a limited number of capsid protein subunits and can increase their volume relatively easily by using the same proteins repeatedly. This limits the number of different capsid protein genes required to form the virion and thus maintains a compact viral genome. Particles made of substructures arranged with icosahedral symmetry are highly efficient ways of forming large-volume three-dimensional structures with repeating subunits. Helical symmetry not only matches nucleic acid symmetry but also allows capsid protein subunits to have identical interactions with each other, except at their ends (Crick and Watson, 1956). The geometrical efficiency of virus capsids is thought to be due to the fact that nucleic acid is relatively inefficient for encoding proteins (three nucleotides required for each amino acid) and so is relatively large. Therefore, using one or a few gene products multiple times and arranging them symmetrically is an optimal strategy for limiting the genome size. Interestingly, some of the viruses with the largest genomes also have icosahedral virions (Wilson et al., 2009; Fischer et al., 2010; Aherfi et al., 2016). There are some notable exceptions to this geometric conservation “rule,” particularly those of the “extreme” viruses of Archaea (see Section 2.5 below).

Given that virion structures are unique and distinctive, it is tempting to use virions as biosignatures. However, to visualize most virion morphologies, a transmission electron microscope (TEM) is required. While this may not be difficult on Earth, it seems unlikely that a TEM will be put on a spacecraft in the foreseeable future. There has been progress with scanning electron microscope (SEM) technology for space exploration, but currently this SEM does not have high-enough resolution to identify virion structures (Edmunson et al., 2016). There has also been progress in enumerating virus-sized particles using dynamic light scattering, nanoparticle tracking analysis, and promising nanopore technology (Schmidt and Hawkins, 2016; Yang and Yamamoto, 2016), but, to our knowledge, none of these techniques have been applied to natural samples containing unknown virions. Moreover, these techniques provide minimal shape information, limiting their use for definitive virion identification. However, these techniques could be useful to pre-sort virus-sized particles for downstream processing possibly including sequencing. They are also more adaptable to mission instrumentation than a TEM.

2.5. Astonishing Virion Diversity in Extremophiles

Viruses of bacteria are dominated by so called head-tail bacteriophages (Fig. 2J). They are taxonomically classified into three viral families, Myoviridae, Podoviridae, and Siphoviridae (Fig. 3), differentiated by their tail structures, and together form the order Caudovirales (King et al., 2011). These head-tail viruses have an icosahedral “head” filled with a double-stranded DNA (dsDNA) genome, connected to a linear cylindrical “tail” which is used to attach and inject the viral genome into the host bacterium.
FIG. 1. (A) Schematic virus replication cycle. Upper left: virion inserts the viral genome into a host cell. In many cases (left arrows), the virus genome is replicated and packaged in newly formed virions, often lysing the host cell in the process. In other cases (right arrows), the virus genome persists in the host cell during cellular replication. Sometimes the viral genome is integrated into the host genome (circle), forming a provirus or prophage. The integrated virus genome replicates with the cellular replication cycle. The provirus (prophage) can be induced to produce new virions, often entering the lytic cycle. (B) The Baltimore classification of viruses by packaged nucleic acid type and route for producing RNA (Baltimore, 1971). The type of nucleic acid genome packaged in virions, RNA or DNA, and single stranded (ss) or double stranded (ds), is indicated in upper ovals. Positive (+) strand refers to the coding sequence of the mRNA. Negative (−) refers to the complement of the mRNA. Nucleic acid intermediates used to form mRNA are shown in lower ovals. Group I comprises dsDNA viruses; group II comprises ssDNA viruses; group III, dsRNA viruses; group IV, positively stranded (+) ssRNA viruses; group V, negatively stranded (−) ssRNA viruses. These five groups correspond to non-retroviral viruses. The discovery of retroviruses (RT) that replicate their genomes through both DNA and RNA intermediates led to the establishment of group VI, which includes (+) ssRNA-RT viruses, and group VII, which includes dsDNA-RT viruses which package ssRNA or dsDNA in their virions, respectively.
upon infection. Direct TEM observation of virus-like particles (VLPs) from soil or aquatic environmental samples shows that over 95% of these nano-sized (>500 nm in diameter) particles have the head-tail morphology (Ackermann and Prangishvili, 2012). A few bacterial viruses with other virion structures are known, such as filamentous or tailless polyhedral structures. Bacterial viruses are presently classified into nine taxonomic families (King et al., 2011). Most eukaryotic virions contain some variation of the icosahedral (Fig. 2A) or helical (Fig. 2D) morphology, and many are enveloped with rounded or pleomorphic morphology (Fig. 2E, 2K, 2M).

On the other hand, archaeal viruses have surprisingly diverse morphology. Some archaeal viral morphotypes are found in bacterial or eukaryal viruses, such as polyhedral or icosahedral (Turriviridae, Sphaerolipoviridae) (Fig. 2A) (Rice et al., 2004; Bamford et al., 2005), helical (Rudiviridae, Lipothrixviridae, Clavaviridae, and recently proposed “Tristromaviridae”) (Fig. 2D) (Janevovic et al., 1983; Zillig et al., 1993; Arnold et al., 2000a; Mochizuki et al., 2010; Rensen et al., 2016), globular/spherical (Globuloviridae) (Fig. 2C) (Häring et al., 2004), and pleomorphic (Pleolipoviridae) (Fig. 2M) (Pietilä et al., 2009). In addition, many archaeal viruses also display morphologies that were never observed previously. These include lemon- or spindle-shaped (Fuselloviridae) (Fig. 2B) (Martin et al., 1984; Wood et al., 1989; Bath and Dyall-Smith, 1998), tailed spindle-shaped (Bicaudaviridae) (Häring et al., 2005a), droplet-shaped (Gutaviridae) (Fig. 2B) (Arnold et al., 2000b), bottle-shaped (Ampullaviridae) (Fig. 2I) (Häring et al., 2005b), and coil- or spring-shaped (Spiraviridae) (Fig. 2N) (Mochizuki et al., 2012) viruses. Together, despite the much smaller number of isolates, viruses of the Archaea are currently classified into 15 ICTV-approved families—well exceeding the nine bacterial virus families (Fig. 3). Environments where Archaea dominate, such as hot springs (Rice et al., 2001; Prangishvili, 2003; Häring et al., 2005b; Bize et al., 2008; Mochizuki et al., 2010) or salt lakes (Sime-Ngando et al., 2011) are filled with remarkably diverse VLPs. The Archaea-specific lemon- or spindle-shaped viruses (Fig. 2B) are frequently isolated from similar Archaea-rich environments, from acidic (Martin et al., 1984; Prangishvili, 2003) and neutral hot springs (Mochizuki et al., 2010), deep sea hydrothermal vents (Geslin et al., 2003; Gorlas et al., 2012) and salt lakes, and isolated from both crenarchaeal and euryarchaeal hosts. Archaeal viruses with virions similar to the classical head-tail bacterial viruses (Fig. 2J) have also been observed and isolated, mainly from euryarchaeal halophiles (Atanasova et al., 2012), but no members of the Caudovirales have been isolated from thermophilic archaea. However, Caudovirales have been found in thermophilic bacteria, infecting hosts such as Thermus growing around 60–80°C (Yu et al., 2006). Any astrovirological virion detection method must encompass this diversity.

2.6. Even Giant Viruses Have Icosahedral Virions

The “small” size of viruses, more specifically virions, has been challenged with the discovery of the so-called giant viruses, also known as “giruses” or Megavirales. This started with the startling discovery of the Mimivirus that infects amoeba in 2003 (La Scola et al., 2003). The Mimivirus has a virion and genome as large as some bacteria. Due

![FIG. 2. Generalized representations of virion morphologies: (A) icosahedral, (B) lemon- or spindle-shaped, (C) spherical, (D) filamentous or helical or rod-shaped, (E) prolate ellipsoid with or without envelope (dashed line), (F) twinned icosahedra, (G) occluded, (H) budded, (I) bottle-shaped, (J) head-tail, (K) ovoid, (L) allantoid, (M) pleomorphic, (N) spring- or coil-shaped.](image-url)
to its large size, Mimivirus was mistakenly identified as a bacterium, "Bradfordococcus," for many years, and only later determined to be a virus. This virus was named Mimivirus because some of its features, including its size, "mimicked" bacteria, yet most of these giant viruses still have icosahedrally symmetric virions. The discovery of giant viruses continues, with the recent findings of Pandoravirus with a 1 μm virion and 1.9–2.5 million base pair (bp) genome and Pithovirus with a virion over 1.5 μm and a 600 thousand bp genome (Legendre et al., 2014). Surprisingly, the genome of the Pandoravirus and the virion of the Pithovirus are larger than many prokaryotic cells and even some small eukaryotes (Philippe et al., 2013). Some giant viruses are parasitized by other viruses, the virophage. The first of these to be discovered

FIG. 3. Virus taxonomy: In order to provide a uniform virus nomenclature, viruses have been classified taxonomically since the 1970s by the International Committee on Taxonomy of Viruses (ICTV) (King et al., 2011). The fundamental ICTV definition of a virus species is "a polythetic class of viruses that constitute a replicating lineage and occupy a particular ecological niche where the polythetic class is one whose members have several properties in common." Virus species differ from the higher viral taxa, which are universal classes defined by strict inclusion requirements. The ICTV classifies viruses in species, genus, (subfamily), family, and in some cases order (King et al., 2011). It is not clear how closely ICTV taxonomic classification reflects phylogeny, particularly given rampant recombination between genes in different viruses. This figure represents the ICTV taxonomic grouping of viruses with viral families and subfamilies, and they are grouped by their packaged nucleic acid and replication as shown in Fig. 1B.
was Sputnik (La Scola et al., 2008) Other virophage are closely related to and may be derived from transposable elements (Fischer and Hackl, 2016). The presence of virophage and their relationships to capsidless transposable elements further blurs the distinction between viruses and capsidless parasitic replicators (Koonin and Dolja, 2014).

2.7. Viruses Are Astronomically Abundant, Particularly in Terrestrial Oceans

Regardless of the debate whether or not viruses are to be considered alive, viruses are critical for structuring Earth’s biosphere. They are extremely abundant, diverse, and play important roles in global biogeochemical cycles. Only in the last few decades have scientists realized that viruses outnumber cellular life-forms on our planet by at least 10 fold, likely even more (Suttle, 2005). A number of researchers have shown, based on nucleic acid content or TEM enumeration, that in 1 mL of seawater, in general, there are from 1–10 million viral particles (Bergh et al., 1989; Proctor and Fuhrman, 1990; Brum et al., 2013). Viruses have been found in all environments where life has been found on Earth, for example, the Arctic sea ice (Wells and Deming, 2006), boiling acid hot springs (Rachel et al., 2002; Prangishvili, 2013), deep sea hydrothermal vents (Geslin et al., 2003) and the submarine mantle (Nigro et al., 2017). All cells are expected to have infectious viruses associated with them. The latest estimation is that one bacterial cell can be infected by 10 viral species on average. Based on these numbers, estimated total numbers of virions on planet Earth are truly astronomical, with a calculated 10^{31} virions in the oceans alone (Rohwer, 2003; Suttle, 2005). Given the extremely large numbers of viral particles in Earth’s oceans, it would be very interesting to determine if there are virions in recently observed water plumes from Europa (Sparks et al., 2016) and previously observed plumes from Enceladus (Hansen et al., 2008). Unfortunately, to our knowledge, no mission is currently planned that will screen for extraterrestrial viral particles in these plumes. The techniques to be used to screen these plumes for virions are unclear at this time, but nanopore and nanoparticle imaging or novel microscopy techniques may be applicable to this endeavor.

2.8. Virus Genomes—The Dark Matter of Sequence Space

Viruses are unique in the current biosphere in that they use either DNA or RNA in single- or double-stranded forms for their genetic material. Without exception, cellular genomes are exclusively composed of dsDNA. Some viruses only use RNA for their entire replication cycle, leading some researchers to postulate that they are remnants of, or direct descendants from, the hypothetical “RNA world” (Koonin et al., 2006). Virus genomes can range in size from less than 2000 bases of single-stranded RNA (ssDNA) in some circoviruses to almost 2.5 million bp of double-stranded DNA (dsDNA) in Pandoravirus (Aherfi et al., 2016). The smallest virus genomes encode only two genes, a capsid protein gene and a protein required for virus genome replication. By contrast, some of the largest virus genomes encode genes previously thought to be only present in cellular genomes, for example aminoacyl tRNA synthetases (Raoult et al., 2004). However, no virus genome isolated to date contains a ribosomal subunit protein or ribosomal RNA gene (Schulz et al., 2017). All viruses are thus absolutely dependent on cellular translation machinery to produce virus proteins. This ribosome dependence led to a recently suggested viro-centric definition of cells and viruses; cells are ribosome-encoding organisms, whereas viruses are capsid-encoding organisms (Raoult and Forterre, 2008; Forterre, 2010).

Plants are known to harbor many so-called viroids, which are small structured pathogenic RNA molecules that are made by cellular DNA-dependent RNA polymerase enzymes that do not encode capsids. Current viroids are unlikely to be derived from ancestral RNA parasitic replicators, as they are confined to plants and depend on cellular RNA polymerases, but they could serve as models for archaic parasitic RNA replicators (Koonin and Dolja, 2014).

Many viral genomes contain genes that have no homologs except in closely related viruses. When a viral genome of an unknown group (i.e., a member of a new viral family) is sequenced, functions can be annotated for a very few genes, and often far more than 50% of the putative genes can only be described as “hypothetical protein” (Pedulla et al., 2003; Mochizuki et al., 2010; Pope et al., 2015).

The vast majority of unknown viral sequence is assumed to be of viral origin, but the lack of corresponding reference sequences in the public database does not allow definitive determination. These observations have led a number of researchers to make the analogy between these putative viral sequences and the presence of dark matter or dark energy in the Universe, by calling these sequences “biological dark matter” (Pedulla et al., 2003) or “the dark matter of the biosphere” (Fileé et al., 2005; Hatfull, 2015).

To date, all isolated viruses of the Archaea have DNA genomes often containing more than 80% of putative genes lacking homologs in the sequence databases (Snyder et al., 2015; Prangishvili et al., 2016). For a long time, archaeal viruses were only found to contain dsDNA genomes, and it was speculated that, owing to the extreme physicochemical conditions under which many archaea thrive, the only possible genome type is dsDNA. However, this view has been overturned with ssDNA archaeal viruses found in halophilic archaea (Pleolipoviridae) (Pietilä et al., 2009), and in hyperthermophilic archaea (Spiraviridae) (Mochizuki et al., 2012). The hyperthermophilic spiravirus Aeropyrum coil-shaped virus (ACV) has the largest known genome of all ssDNA viruses, with 24 thousand bases, which is more than double the size of the previously largest known ssDNA genome. This is particularly surprising given that ACV was isolated from a host organism that grows at 90°C (Mochizuki et al., 2012).

2.9. Virus Metagenomes

The recent development of high-throughput DNA sequencing has accelerated studies on genomic diversity of environmental samples in a culture-independent manner, known as metagenomics. However, unlike for cellular samples, where 95% of sequences are similar to previously known sequences, metagenomic analysis targeting virus-sized particles (metavirionomes) in many natural environments generates 60–90% of sequences that do not have clear homologs (Edwards and Rohwer, 2005; Rosario and Breitbart, 2011; Hurwitz and Sullivan, 2013; Yoshida et al., 2013; Simmonds et al., 2017). For a number of reasons, marine viromes have been among the most intensively
studied. Early metavirome studies first indicated an overwhelming diversity of virus sequences with possibly billions of virus species (Rohwer, 2003). More recent studies indicate that there may only be up to a million virus species (Brum et al., 2015). Nonetheless, there are still many new groups of viruses that are known only by their sequence (Breitbart, 2012; Brum et al., 2015; Roux et al., 2016). A recent analysis of all published viral DNA metagenomes found that nearly 70% of virus-associated genes were novel (Paez-Espino et al., 2016). Thus there is still a great deal to discover in our virus world.

2.10. CRISPR—A Virus Defense Mechanism and Historical Record

Most archaeal and some bacterial cellular genomes contain so-called CRISPRs (Clustered Regularly Interspersed Palindromic Repeats), a previously cryptic sequence structure that is now known to provide acquired immunity to virus infection (reviewed in Koonin, 2017). CRISPR sequences consist of repeated DNA sequences with short so-called spacer sequences between the repeats that are identical to parts of virus genome sequences or other mobile DNA (Koonin, 2017). One of these systems, the CRISPR/Cas9 system, has been widely used for genome engineering (Wright et al., 2016). However, more importantly for astrobiology, CRISPR sequences also contain a record of past virus infections. This has proved extremely useful for determining past infections (Wilmes et al., 2009), reconstructing virus genomes (Anderson and Banfield, 2008), inferring hosts for unknown viruses (Anderson et al., 2011), and virus biogeography (Held and Whitaker, 2009).

2.11. Viruses Change Host Metabolism, Sometimes to the Host’s Advantage

In many known cases virus infection is deleterious to the host cell or the organism. However, more recently, it has been realized that some, if not most, viruses coexist symbiotically with their hosts (reviewed in Roossinck, 2011; Metcalf and Bordenstein, 2012; Roossinck, 2015). Interactions between viruses and their hosts can range from parasitism to mutualism (e.g., Hamelin et al., 2017). There are a growing number of examples of virus infections that clearly benefit their hosts, mostly by contributing genes that allow their hosts to survive under conditions under which they would otherwise perish (reviewed in Roossinck, 2011; Roossinck, 2016; Roux et al., 2016).

Upon viral infection, virus-encoded genes are expressed. Most of these genes are specific for virus replication. However, in many cases, viral genomes also encode genes whose products change their host’s metabolism. The best-characterized of these genes are components of photosystem II, the machinery for oxygenic photosynthesis. These genes are found in viruses that infect photosynthetic cyanobacteria and not only allow the infected cyanobacterium to undergo photosynthesis under conditions that uninfected bacteria could not but also enable them to thrive in parts of the ocean that uninfected cells cannot (Mann et al., 2003; Lindell et al., 2005; Sullivan et al., 2006). Moreover, other so-called auxiliary metabolic genes (AMG), important for nucleotide, phosphate, carbon, sulfur, nitrogen, and ammonium metabolism, have been reported in virus genomes and metagenomes (Breitbart, 2012; Anantharaman et al., 2014; Roux et al., 2016). It would be very interesting to determine if these virus-encoded gene products lead to isotopic fractionation that is different from cellular enzymes, which could lead to the development of virus-specific isotopic biomarkers. It is probable that extraterrestrial viruses—if they exist—will also coexist with host cells on a spectrum between mutualism and parasitism.

3.1. Viruses Are Critical for Biogeochemical Cycles

Due to their extremely large numbers, in addition to allowing hosts to adapt to new environments and changing their metabolism, viruses also contribute to global nutrient cycling. Viruses are assumed to cause 50% of bacterial death in the modern oceans, with a calculated $10^{28}$ virus infections/day (Suttle, 2007; Breitbart, 2012). Each time a viral infection causes cell lysis, cellular organic matter is released to the surrounding environment. These burst cells provide nutrients for other organisms and maintain the food chain, via the so-called viral shunt (Wommack and Colwell, 2000; Weinbauer, 2004). Up to 25% of fixed carbon is predicted to be recycled in the upper ocean water column due to the activity of these marine viruses (Wilhelm and Suttle, 1999). If there were no virus-mediated cell lysis, cells would gradually fall to the seafloor, thereby not only removing nutrients from the upper ocean but also lowering the carbon content of the oceans. This process of carbon recycling is an essential component of global climate regulation (Suttle, 2007). Viral lysis of abundant coccolithophore algae may influence climate by release of cloud-nucleating chemicals (Wilson et al., 2009). Moreover, about 5% of the oxygen on Earth is estimated to be produced by virus-infected cells using virally encoded photosynthesis genes (Suttle, 2007). It is intriguing to speculate on the role of viruses in “the great oxygenation event” and the amount of oxygen and organic matter in the ancient oceans (Holland, 2006). To our knowledge viruses have not been incorporated in such models. Microbial ecologists are beginning to include viruses in terran biogeochemical cycle models (Stec et al., 2017), and we suggest that modelers of exoplanets and their satellites also consider including viruses in their models.

3.2. Viruses Drive Evolution

It has been said that evolution and possibly life as we know it would not be possible without viruses (Greene and Reid, 2014). There are two major ways that viruses drive evolution: first, by providing strong selective pressure for resistance to viral infection; and second, by moving exogenous genetic materials into cells. Evolution of host resistance to viral infection, particularly against a virus that kills its host, provides exceptionally strong evolutionary pressure. On the other hand, since viruses require hosts for their replication, there is also strong evolutionary pressure on the virus to overcome host defenses. This interplay of host and viral evolution has been demonstrated both experimentally and theoretically (Turner and Chao, 1999; Wichman and Brown, 2010; Elena, 2016). Viruses and their hosts are thus often said to be in an “arms race” or “running to stand still,” fulfilling the “Red Queen hypothesis” of evolution (Van Valen, 1973). In simulations of early life, parasitic replicators, similar in concept to what we now know as viruses, inevitably emerge (Koonin, 2011).
Less appreciated but probably no less important for evolution is that viruses drive evolution by the introduction of new genes to cells and by transferring genes from one cell to another. This acquisition of genes from a source other than a direct ancestor is termed horizontal gene transfer (HGT). Viruses are known to be major agents, if not the most important agent, of HGT. Horizontally transferred genes are often acquired by the host genome, particularly if they confer the recipient with an evolutionary advantage. Since genetic diversity in the viral gene pool is very high and viruses are agents of HGT, the successful transfer of novel viral genes into host genomes is likely commonplace. Virus-mediated HGT is particularly important for evolution in species that reproduce asexually but is also widespread in both vertebrates and invertebrate animals (Koonin, 2016). For example, the human genome contains between 8% and 40% virus-derived sequences, some of which have been beneficial for human development (Lander et al., 2001; Feschotte and Gilbert, 2012). One well-studied example is the syncytin gene that is essential for mammalian placental development. Syncytin is derived from not one but multiple apparently independent ancient retrovirus infections (Lavialle et al., 2013).

Some viruses insert their genomes into their host’s genomes (Fig. 1A). The best-studied of these are the retroviruses that require insertion of their genome into their host genome as part of their replication cycle. This integration step generates a copy of the virus genome within the host genome, in the form termed a provirus. Under certain conditions, proviruses can be triggered to replicate and produce new virions. As their proviruses replicate, retroviruses can acquire cellular genes. In fact this is how the first oncogenes (cellular genes that cause cancer when misregulated) were discovered (Stehelin et al., 1976). Many bacterial and archaeal DNA viruses also integrate their genomes into the host genome, generating proviruses (or prophages, in the case of bacterial viruses). Upon replication, these prophages can also incorporate host genes and transfer them to other cells in a process called transduction. Some viruses and defective viruses, known as generalized transducing bacteriophage and gene transfer agents (GTAs) respectively, incorporate random DNA into their virions and inject these DNAs into host cells, leading to massive amounts of HGT. Generalized transducing bacteriophage have been widely used in bacterial genetics (Bachmann, 1990) and seem to be important for the development of both bacterial pathogenesis and antibiotic resistance (Moon et al., 2015). Even viruses that use ssDNA or RNA as their genetic material are known to be incorporated, at least partially, into cellular dsDNA genomes (Stedman, 2015). Due to transduction and promiscuous provirus gene acquisition, the amount of horizontal gene flow via viral infection is currently considered much higher than once expected (McDaniel et al., 2010) and is hypothesized to have been even higher in primordial environments (Woese, 2002; Koonin et al., 2006). It is thus highly likely that viruses played a major role in the evolution of early life.

3.3. Viruses as Drivers of Major Evolutionary Transitions

Viruses have been hypothesized to be responsible for the evolution of DNA as the genetic material, the three domains of cellular life, the development of the eukaryotic nucleus, and even multicellularity. One of the first proposed virus-driven evolutionary events is that viruses may have “invented” DNA genomes in an ancestral RNA-protein world, and that DNA was adopted later by cellular organisms (Villarreal and DeFilippis, 2000; Forterre, 2002, 2005). In order to convert RNA into DNA, three steps are required: a ribonucleotide reductase that transforms ribose to deoxyribose, a thymidylate kinase that adds a phosphate group to thymidine, and a reverse-transcriptase that makes a DNA copy from an RNA template. Interestingly, all of these genes have been found in virus genomes, and these viral genes have much greater sequence diversity than the corresponding cellular genes. The most parsimonious explanation for this diversity is that these genes first arose in viruses, evolved therein, and some genes were transmitted to, and retained by, cellular genomes. Moreover, phylogenetic analysis of DNA polymerase genes often places viral genes at the base of the eukaryotic clade, indicating that the viral polymerases are ancestral (Filée et al., 2002; Koonin et al., 2006; Villarreal and DeFilippis, 2000).

An extended version of this hypothesis is that different virus infections led to the formation of the modern three cellular domains of life (Forterre, 2006a). The argument is that DNA polymerases in the three cellular domains are very different from each other, but each is clearly related to extant viruses, which also have orthogonal DNA replication systems (Woese, 2002). Engulfment of a large DNA virus has also been proposed as the origin of the eukaryotic nucleus (Bell, 2001; Takemura, 2001). The recent discovery of the giant Mimiviruses has caused these theories to be revisited (Bell, 2009). The intracellular compartment formed in the Mimivirus’ host amoeba, termed the virus factory (or virion factory), closely resembles the eukaryotic nucleus in both structure and function (Raoult and Forterre, 2008). Thus an ancestral virus infection that caused the formation of a virus factory may have led to the development of the nucleus. A modern infection of a bacteria with a viral virus leads to the development of a remarkably nucleus-like structure in the infected bacterium (Chaikertasik et al., 2017).

The remarkable morphological and sequence diversity observed in archaeal viruses provides another clue regarding cellular (and virus) evolution. A naive assumption would be that the archaeal virosphere would have limited diversity, due to the harsh living conditions of most archaeal virus hosts. But the complete opposite, high diversity observed in archaeal viruses and limited diversity in bacterial viruses, is indeed puzzling (Rachel et al., 2002; Ackermann and Prangishvili, 2012; Prangishvili et al., 2016). One recently proposed hypothesis incorporates ancient viruses as a direct evolutionary driving force, which eventually led to the divergence of the Bacteria and Archaea (Forterre and Prangishvili, 2009: Prangishvili, 2013). They suggest that ancient viruses were already rich in morphological diversity at the time of the Last Universal Common Ancestor (LUCA). Later, one group of cellular organisms, the ancestors of bacteria, developed a cell wall to protect themselves from viral infections. In order to overcome this rigid defense system, bacterial viruses evolved a physical penetration mechanism to inject their genetic material, the tail, which led to the predominance of head-tail viruses in the bacterial domain. On the other hand, other primordial cells, the ancestors of the Archaea, altered their cell surface proteins by
glycosylation to escape from viral recognition. Their viruses then glycosylated their capsid proteins in order to attach to the hosts. This explains the high level of glycosylation on archaeal cell surface proteins and archaeal virus capsid proteins; such glycosylation is not frequently observed in bacterial cells and viruses.

Finally, viruses may have been involved in the evolution of multicellularity. In the presence of excess resources, multicellularity could provide protection from virus infection, as cells on the outside of an aggregate would be more susceptible to virus infection than those inside (Ruardij et al., 2005; Greene and Reid, 2014). Moreover, the “invention” of apoptosis or programmed cell death, which is directly coupled to the evolution of multicellularity, could be due to virus infection pressures (Iranzo et al., 2014).

3.4. Diversity of Viruses and Their Evolutionary Relationships

Due to the lack of “universal” genes, such as 16/18S rRNA genes, in viruses, it is nearly impossible to place all viruses into one phylogeny to represent their evolutionary relationships. Nucleotide polymerases are often (but not necessarily always) encoded in virus genomes, and they have been used frequently to reveal the history of some distinctly related viruses (Koonin, 1991; Koonin et al., 2008; Koonin and Dolja, 2014). But the polymerases themselves are highly diverse, both in terms of sequence and substrates used. Most viruses encode a nucleotide polymerase, but some, including most hyperthermophilic archaeal viruses, do not encode a recognizable polymerase gene in their genomes.

However, more recently, comparison of the three-dimensional structure of major capsid proteins indicates that several viruses that infect archaea, bacteria, and eukarya have a common double beta-barrel capsid protein structure and are thus extremely distantly related, even though sequence similarity cannot be detected (Rice et al., 2004; Khayat et al., 2005; Krupović and Bamford, 2008). Similar results were obtained with a different capsid protein structure found in a haloarchaeal virus, bacterial viruses, and herpesviruses (Pietilä et al., 2013). Thus classification of different viral groups based on capsid protein structure seems to be an effective way to discern phylogenetic relationships between distinctly related viruses infecting viruses of all three cellular domains (Krupović and Bamford, 2011; Abrescia et al., 2012). Major capsid protein structures could potentially become the viral equivalent to the cellular 16S/18S rRNA gene, especially since all viruses encode capsid proteins. However, identification of the capsid protein gene is not straightforward, particularly when analyzing new viral groups from metagenomic analysis (Mokili et al., 2012) and determining capsid protein structures is highly problematic in astrobiological samples.

However, even if capsid protein structures could provide a comprehensive virus phylogeny, there appears to be rampant recombination in virus genomes, such that many of the genes in a virus genome do not have the same phylogeny. An extreme example is the chimeric viruses that appear to have undergone recombination between RNA (Baltimore Class IV) and DNA (Baltimore Class II) viruses (Diemer and Stedman, 2012; Roux et al., 2013). A recent metagenomic survey of invertebrate animals discovered hundreds of new RNA viruses, indicated that recombination between different viral genes is rampant, and confounds virus phylogeny (Shi et al., 2016). Thus it may be impossible to represent virus phylogeny by a tree-like structure, as viruses probably have a more mosaic or modular genome evolution dominated by HGT (Botstein, 1980; Hendrix et al., 1999; Koonin and Dolja, 2014; Koonin et al., 2015; Iranzo et al., 2016).

3.5. What Is the Origin of Viruses?

The evolution of cellular and viral life may not easily be separable (Szathmáry and Smith, 1997; Villarreal, 2004; Koonin et al., 2006; Jalasvuori and Bamford, 2008; Forrester, 2010; Koonin and Dolja, 2013). There are three predominant theories for virus origins: First is that viruses are escaped cellular genes. Second is that viruses are degenerated cells. Both of these theories indicate a relatively recent origin for viruses. The third theory is that viruses could have an independent origin or multiple independent origins (Koonin et al., 2006). Increasingly, virus researchers think that the latter is true (Forrester, 2006b; Koonin et al., 2006; Witzany, 2008; Forrester and Prangishvili, 2009; Forrester and Krupović, 2012; Nasir and Caetano-Anolles, 2015). There are also a few researchers who controversially claim that viruses are remnants of an otherwise extinct fourth domain of cellular life (Philippe et al., 2013), but see the works of Schulz et al. (2017) and Moreira and López-García (2015). Nevertheless, if viruses were involved in the origin of DNA, the transition between an RNA and DNA world, the origin of the three domains of life, the development of the bacterial cell wall, or the origin of the eukaryotic nucleus, they must be very ancient.

One piece of evidence is that viruses are the only extant terran life-forms that use RNA genomes, and could thus be descendants of viruses or similar replicative entities that existed in the hypothetical RNA world (Sankaran, 2016). An active area of origins-of-life research is the encapsidation of self-replicating RNAs in vesicles, not unlike enveloped RNA viruses (Robertson and Joyce, 2014; Deamer, 2017). There are also many genes that are found in virus genomes but not in any cellular genomes, further supporting the independent origin hypothesis (Koonin et al., 2006; Krupović and Bamford, 2010). One of the great mysteries of the origin of viruses is the origin of the defining characteristic of viruses, the vi- rion. Recent analyses indicate that there may have been multiple origins of capsids and thus multiple virus origins (Krupović and Koonin, 2017). Unfortunately, there is no direct evidence for ancient viruses in the geological record, but there is considerable indirect evidence for virus antiquity.

3.6. What Is the Evidence for Ancient Viruses?

Indirect evidence for ancient viruses comes from genomic and structural studies. The presence of common viral genes in genomes of organisms known to have diverged tens of millions of years ago from both fossil and genomic data, together with common antiviral genes, indicates that specific viruses and viral defenses are at least tens to hundreds of millions of years old (Emerman and Malik, 2010; Aiewsakun and Katzourakis, 2017). Unfortunately, sequence similarity becomes undetectable for more ancient divergences. The presence of conserved capsid protein structures is
indirect evidence for ancient viruses’ lineages (see Section 3.4 above). Moreover, when virus capsids are classified based on structures of their capsid proteins, there are a limited number of clades. The most parsimonious explanation for these data is that these virion protein structures predated the divergence of the three domains of cellular life (Rice et al., 2004; Krupovič and Bamford, 2011). However, there may have been multiple, independent acquisitions of virus capsid proteins throughout both viral and cellular evolution (Krupovič and Koonin, 2017).

There is almost no evidence for viruses in the rock record. Structures resembling virus inclusions formed by current insect viruses have been detected in ca. 100-million-year-old amber (Poinar and Poinar, 2005). Laboratory silicification of the well-studied bacteriophage T4 and extremophile viruses led to the determination that the spatial resolution of current techniques is insufficient to determine whether organic inclusions in rock are viral in origin (Laidler and Stedman, 2010; Orange et al., 2011). In both analyses, detectable virion morphology was lost over periods of days to months (Laidler and Stedman, 2010; Orange et al., 2011).

### 3.7. Could Viruses Be Spread Extraterrestrially?

At present there is no evidence for virions in any extraterrestrial samples, mainly due to lack of biomarkers. However, considering the high abundance in the modern biosphere and due to the small particle size and mass of most virions, it is likely that meteoritic bombardment or possibly volcanic activity would have ejected viral particles from Earth. A few studies have examined virus stability in the space environment. Tobacco mosaic virus, a plant virus that is highly resistant to desiccation and can be crystallized, was exposed to simulated interstellar radiation and low temperature (77 K) high vacuum (10⁻⁸ torr) with minimal infectivity loss (Koike et al., 1992). Poliovirus and bacteriophage T1 were exposed to the space environment on high-altitude balloons and rockets for short times without significant loss of activity (Parvenov and Lukin, 1973). Dried bacteriophage T7 was used as a biological dosimeter for UV irradiation on the International Space Station (ISS), but infectivity was not tested thereafter (Bérczes et al., 2015). A survey of virion stability studies was published (Griffin, 2013), but none of these studies addresses possible survival of viruses to interplanetary transit. It is theoretically possible that endolithic viruses or viruses encased in salt crystals could survive space exposure, but for how long is not clear. We found that silica-coated bacteriophage T4 could survive a month of drying, but infectivity was completely lost thereafter (Laidler et al., 2013). Samples have just been returned from the Tanpopo mission on the ISS (Kawaguchi et al., 2016). The Tanpopo mission collected particles present in low Earth orbit on the ISS in aerogel for over a year and also exposed terrestrial bacteria and yeast to the space environment for up to 2 years. Whether any micro-meteoritic samples collected in the Tanpopo aerogel contain virions would be very interesting but unlikely to be unambiguous due to the lack of validated virus biosignatures.

### 3.8. Virus Biosignature Detection

What might virus biosignatures be? The 2015 NASA Astrobiology Strategy describes the breakdown of current techniques and strategies for life detection into remote and direct detection of extinct and existing life. Although viruses have been shown to impact life at an ecosystem level, remote detection of individual virions is implausible; however, macroscopic effects of virus infection could be detectable, given changes to Earth’s biogeochemical cycles (see Section 3.1). Direct detection of viruses is more realistic given the average size of a virion.

The standard detection strategies for extinct life, including chemical signatures and isotopic fractionation, also serve as detection strategies for extinct viruses. While the standard array of chemical signatures for detecting extinct life includes hopanoids, sterols, cyclic alkanes, isoprenoids, and carotenoids, specific virus chemical signatures are still under investigation. Enveloped viruses containing lipids may provide a detectable chemical signature (Kyle et al., 2012). Unlike most host-derived viral components, some viral lipids exhibit a distinct chemical signature from the lipids of host organisms; such a distinction may prove critical in the chemical detection of viruses in the rock record. Moreover, some viruses also have what was thought to be virus-specific nucleotide modifications, for example, 5-hydroxy-methyl cytosine in bacteriophage T4. However, this modified nucleotide has recently been found in human DNA (Bachman et al., 2014).

Common techniques for biosignature detection, such as extraction of rocks with organic solvents followed by liquid chromatography and demineralization, can be adapted for use in virus detection by simple manipulations of equipment and technique. Laboratory methods to identify the isotopic abundances (high-resolution mass spectrometry), chemical species (GC-MS, NMR, or FTIR spectroscopy), and elemental composition (inductively coupled argon plasma mass spectrometry) all can be used with little to no modification to identify and characterize extant virus samples along with samples of potential life. However, differentiating viruses from potential cellular material in bulk samples is problematic. More promising is detection of a virus-mediated event, such as massive lysis or death of host organisms. Such an event has been proposed to have led to the famous chalk deposits of the white cliffs of Dover. These deposits appear to be due to massive lysis of coccolithophore algae, possibly due to virus infection (Wilson et al., 2009).

Because virions do not metabolize directly, no by-products such as methane or N₂O should be detectable remotely; however, if host metabolism is changed by virus genes, this could be detectable remotely. Similarly, disequilibria such as redox, chemical, and mineralogical chemical signatures should not be used as direct detection mechanisms for viruses but could be useful as indirect measurements. It is not known to our knowledge if virus-specific AMG fractionate isotopes differently than cellular enzymes, but this work should be of high priority.

Standard detection strategies for existing life, including biomolecules and chiral excess, serve as detection strategies for existing viruses. As opposed to cellular life, fully assembled virions are composed of three of the four standard cellular macromolecules including nucleotides (DNA or RNA), amino acids, proteins, (viral capsids), and in some cases lipids (viral envelopes). Virions are smaller than their host; thus the proportional relationship between components in detectable biomolecules of viruses could provide a virus biosignature if high-enough (<10 nm) spatial resolution
could be obtained. Many viruses produce complex capsid geometries from the arrangement of repeated proteins; in some cases, a single protein is sufficient for assembly of a complete capsid (Acheson, 2011). Detection of repeated proteins could provide insight into the sample identity. Sequence identification of the amino acids or nucleotides of a virion provide immediate insight into identity of the virus. In addition to the common techniques applied to extinct viruses described above, techniques such as melting and filtration of ice followed by ultrasensitive analysis using various chemical techniques can be applied for the detection and characterization of known existing viruses. These include, but are not limited to, PCR and phylogenetic analyses, immunoassays, lab-on-a-chip methods using capillary electrophoresis and laser-induced fluorescence, and ultra-sensitive mass spectrometry. These modern techniques, such as high throughput sequencing or high density microarrays (ViroChip), can identify known viruses (e.g., Yozwiak et al., 2012), and some bioinformatics techniques can identify unknown but prevalent viruses (Roux et al., 2016). However, the identification of truly novel viruses, such as potential extraterrestrial viruses, remains problematic. Nonetheless, we find that the development of virus-specific biomarkers should be an active area of astrovirology research and is highly relevant to the detection of viruses in extraterrestrial material.

4.1. Future Outlook

More than a century has passed since the discovery of the first viruses. Entering the second century of virology, we can finally start focusing beyond our own planet. Priorities for future astrovirology research in the short term should include (1) validation of virus biosignatures, (2) consideration of virus-detection experiments to be used for missions that sample water plumes from Enceladus and Europa, (3) inclusion of viruses in models for ancient oceans and extraterrestrial systems, (4) determination whether virus-encoded and transferred AMG fractionate isotopes differently than their cellular counterparts.

In the longer term, astrovirology objectives should include (1) more exposure studies of viruses, particularly endolithic viruses, to the space environment; (2) characterization of highly abundant cosmopolitan viruses (Roux et al., 2016; Urayama et al., 2016); (3) greater research into the roles of viruses in the origin and evolution of early life; and (4) more outreach to astrobiologists and the general public regarding the ubiquity and role of viruses in Earth’s ecosystems.

Viruses are an integral, highly abundant yet underappreciated part of life on Earth. They play a critical role in biogeochemical cycles and evolution today and may have been intimately involved in the origin of life. We do need to learn much more about viruses on modern Earth before we look elsewhere, but let’s start looking.

Acknowledgments

The authors acknowledge Autodesk, Inc., employee Joe Lachoff for his assistance in generating an initial version of Fig. 3 and cover art. Author A.B. contributed to this article while employed by Autodesk, Inc. Author T.M. was supported by Kurita Water and Environment Foundation research grant program, award number 16B094, Japan Prize Foundation, and KAKENHI grant from Ministry of Education, Culture, Sports, Science and Technology of Japan, grant numbers 17H05811 and 17H05229. Author K.S. was supported by Portland State University, the National Science Foundation, Grants MCB 1243963 and MCB 0702020, and NASA, Award number NNA11AC01G.

Author Disclosure Statement


References


Address correspondence to: Kenneth M. Stedman
Biology Department
Portland State University
P.O. Box 751
Portland, OR 97207-0751
E-mail: kstedman@pdx.edu
Submitted 10 January 2017
Accepted 20 July 2017

---

**Abbreviations Used**

- **ACV** = *Aeropyrum* coil-shaped virus
- **AMG** = auxiliary metabolic genes
- **CRISPRs** = Clustered Regularly Interspersed Palindromic Repeats
- **dsDNA** = double-stranded DNA
- **HGT** = horizontal gene transfer
- **ICTV** = International Committee on Taxonomy of Viruses
- **ISS** = International Space Station
- **NAI** = NASA Astrobiology Institute
- **NAIViFoG** = NASA Astrobiology Institute Virus Focus Group
- **SEM** = scanning electron microscope
- **ssDNA** = single-stranded DNA
- **TEM** = transmission electron microscope
- **VLPs** = virus-like particles